Global Change Science Requirements

For

Long-Term Archiving

Report of the Workshop,
Oct 28-30, 1998
National Center for Atmospheric Research
Boulder CO

Sponsored by NASA and NOAA, through the
USGCRP Program Office

March 1999

Global Change Science Requirements

For

Long-Term Archiving

Report of the Workshop,
Oct 28-30, 1998
National Center for Atmospheric Research
Boulder CO

Sponsored by NASA and NOAA, through the
USGCRP Program Office

March 1999

Table of Contents

Acknowledgements

This workshop report was assembled by Greg Hunolt of SGT, Inc., with editing by Cara Sucher of the U.S. Global Change Research Program Office. The workshop was co-chaired by Martha Maiden of the National Aeronautics and Space Administration and Tom Karl of the National Oceanic and Atmospheric Administration. We would like to thank the workshop participants and offer our particular appreciation to the scientists who reviewed and contributed to this report: Richard Armstrong, John Bates, Francis Bretherton, Josef Cihlar, Margarita Conkright, Peter Cornillon, Richard Cram, Lon Hood, Bisher Imam, Jack Kaye, Katherine Kidwell, Gwynneth Martin, Mark Parsons, Cora Randall, Gregory Scharfen, Kevin Trenberth, and Victor Zlotnicki.

David Goodrich
U.S. Global Change Research Program

Executive Summary

The USGCRP workshop "Global Change Science Requirements for Long-Term Archiving" was held at the National Center for Atmospheric Research in Boulder, CO on October 28-30, 1998. Participants included scientists with extensive experience as users or producers of large and/or long-term data sets. Discussions centered around the purpose of a Long-Term Archiving program for Earth observation data and derived products, lessons to be learned from current and past experience, and the vision, guiding principles, and essential functions necessary for the success of such a program.

Data sets targeted for the long-term archiving program include NASA Earth Observing System (EOS) data and derived products; data from Next Generation Weather Radar (NEXRAD), Polar Orbiting Operational Environmental Satellite (POES), Geostationary Operational Environmental Satellite (GOES), the future EOS PM to Natioanl Polar Orbiting Environmental Satellite System (NPOESS) bridge platform and NPOESS; other satellite data; *in situ* data; and additional products needed for interpretation and use of the archive.

What is the purpose of the Long-Term Archiving program?

The Long-Term Archiving program provides a critically needed data and information resource for environmental scientists whose research is used in many public policy and business decisions. It also helps scientists gain new knowledge about the Earth system, and provide a legacy upon which future generations can resolve current and impending issues.

Many economic issues will benefit from the scientific use of the Long-Term Archive. Examples include determining how a power plant should be sized for peak loads; if regional water resources are adequate for periods of drought and high demand; the likelihood of crop failure due to bad weather; and the prospects for commercial fishing with major ocean temperature anomalies.

Other important issues that will profit from the existence of a Long-Term Archiving program include environmental concerns such as the future habitability of Earth. Tropospheric warming, possible productivity decline, extreme weather and climate events, melting ice sheets, anthropogenic climate change, and possible stratospheric ozone decline are just a few of the issues that can be better understood with a well-organized Long-Term Archiving program - issues which require analyses of carefully preserved and fully documented long-term data sets.

Lessons learned from experience with existing data sets:

Long-term archiving has proven to be of enormous benefit when data collected from Earth observing platforms and derived environmental products are preserved in ways that facilitate reanalysis and user access. Data archiving makes the generation of new geophysical results possible; however, preserving complete documentation along with the data is of absolute and fundamental importance for proper study. Every example of successful re-analysis of archived data has depended on quality documentation.

The complex science problems associated with climate research require scientists to use combinations of data sets from both satellite and *in situ* sources. Facilitating the use of multiple data sets will help maximize support for researchers.

Life-cycle data management of the Long-Term Archive is also needed, from the initial planning through the development and implementation. It must be an ongoing process - not stopping when an initial implementation is made. Management should involve researchers, data managers, data collectors and primary users, and must ensure that data management implications of system alternatives are fully explored.

Access to archived data is also key factor in securing its preservation. Only through use is the scientific value of data proven and the expense of preserving it justified. Planning and execution of data migration to a system with improved user access can be as important as a new archive medium. Promoting use of the data is a key element of a data preservation strategy.

Essential functions for a long-term archive

Presently, the lack of a focused Long-Term Archiving program has left archiving decisions to be handled case by case, by agencies in operational and research programs whose primary goals may include use of only near-term, short-term, or real-time data. The vision for a Long-Term Archiving program developed at the Workshop is "A continuing program for preservation and responsive supply of reliable and comprehensive data, products, and information about the Earth system for use in building new knowledge to guide public policy and business decisions". Long-Term Archiving program functions are supported by two guiding principles:

1. A Long-Term Archive must be established and operated in the simplest way possible to meet user needs and program goals, and
2. A Long-Term Archive is not only for today's generation of users, but also for the next generation of scientists and citizens whose needs have yet to be expressed but must be provided for.

The Long-Term Archive program must be actively engaged with its user community, including scientists who are data and product contributors and users, observing system managers, large-volume and small-volume private sector users, and data experts.

The Long-Term Archive program must develop procedures and criteria for determining what data is to be included, excluded, and removed from the archive. It should be driven by present science priorities, scientific assessments, general public needs, and national interests.

The Long-Term Archive program must ensure that archived data sets and products are accompanied by complete, comprehensive, and accurate documentation to ensure usefulness to the user; poor documentation destroys the value of a data set or product. Information about the physical location and access paths to the data must be easily available.

Data and documentation received from operational or research sources must be verified, stored, cataloged, and made available as soon as possible to meet user needs. The ability to access data for re-analysis when improvements are made in data-processing algorithms is required.

The data must be preserved and maintained in perpetuity. Integrity checks during the migration of data from one type of media to another must occur on a routine basis to prevent the data from becoming inaccessible or deteriorating beyond repair.

Customer service and technical representatives are required for user support and to ensure that users' access needs are met. Research points of contact are also required. Near-real time data should be accessible within minutes or hours from time of acquisition, while other archived data should be accessible within hours or days of processing.

Conclusion

The economic well being, quality of life, and economic strength of the United States depend partly on public policy, individual citizens, and business decisions, many of which are now guided by scientific research and environmental assessments. This dependence is increasing due to the growing sensitivity of modern life and commercial activity to the potential impacts of climate change and environmental hazards.

Increase in scientific understanding of the Earth system is the result of an active, growing community of research scientists. Their work requires a wide variety of global-scale measurements sustained over long time periods. The extension, cross-calibration and validation of existing data sets requires combining measurements and derived products from multiple sources. This will only be possible if the data, derived products, and full documentation are preserved by the Long-Term Archive program, and each scientist has easy access to the particular combination of data and information services they require.

To make effective use of the enormous investment our country has already made in Earth Observing Systems, we must embrace the purpose, vision and guiding principles of the Workshop by elevating the existing long-term archive activity to the status of a distinct program with a specific plan and dedicated budget.

1.0 Introduction / Background

This report documents the results of the U.S. Global Change Research Program (USGCRP) workshop on "Global Change Science Requirements for Long-Term Archiving" held at the National Center for Atmospheric Research (NCAR) in Boulder, CO on October 28-30, 1998. Participants included scientists with extensive experience as users or producers of large and/or long-term data sets. The report describes the objectives of the workshop, the purpose of a Long-Term Archiving (LTA) program for Earth observation data and derived products, lessons to be learned from current and past experience, and the vision, guiding principles, and essential functions necessary for the success of such a program.

1.1 Workshop on Global Change Science Requirements for Long-Term Archiving

The "Global Change Science Requirements for Long-Term Archiving" workshop was sponsored jointly by the National Aeronautics and Space Administration (NASA) and the National Oceanic and Atmospheric Administration (NOAA), and organized by the USGCRP Coordination Office. The objectives of the workshop were to develop science requirements for optimal long-term archiving of the large amount of atmospheric and ocean data expected from upcoming studies. The LTA program plans to include data from:

1. The NASA Earth Observing System (EOS) platforms (and other related NASA data),
2. The NOAA / Department of Defense (DOD) National Polar-orbiting Operational Environmental Satellite System (NPOESS), and
3. Expanding *in situ* data streams such as NOAA's Next Generation Weather Radar (NEXRAD) system operated by NOAA, DOD, and the Federal Aviation Administration (FAA).

The workshop brought together scientists with extensive experience as users and producers of large and long-term data sets (see Appendix A for a list of participants). They contributed their knowledge of the many scientific problems associated with global climate monitoring and analysis as well as the necessary demands for properly archived data, derived products and supporting documentation. Data managers from NOAA, NASA, NCAR, and the U.S. Geological Survey (USGS) and representatives of NOAA and NASA management participated in the workshop as well. Their goal was to establish a program that could provide the most comprehensive and accessible long-term Earth science data archive possible.

Drawing on lessons learned from archived data usage over the past decade, workshop participants considered the future observing systems and provided recommendations for an improved data archiving and accessibility program. The recommendations are clarified by a vision statement, an explanation of the LTA program's purpose, and descriptions of the essential functions and characteristics of the program.

1.2 Background

Presently, NOAA is the federal agency with statutory responsibility for long-term archiving the nation's atmospheric and oceanographic data. In a 1989 Memorandum of Understanding (MOU) (see Appendix B), NASA and NOAA agreed that NOAA would eventually assume similar responsibilities for the NASA EOS and other related atmospheric and oceanographic data. The MOU is a high-level agreement laying out general responsibilities of the two agencies, and calls for implementation agreements to address actual plans for accomplishing the transfer of responsibility. NASA and NOAA agreed to begin work on implementing the MOU. In August 1998 they began to develop a plan and secure funding for archiving NASA's atmosphere and ocean data. A critical step in their approach is defining the science requirements for the LTA program that will provide support to the USGCRP, NASA's Earth science research program, and other related research efforts; the primary objective of the LTA program is for use in scientific research and applications. Secondary objectives of the LTA program include ensuring data availability for other needs and data preservation as Federal records within the authority of the National Archive and Records Administration (NARA).

This report will be used as the basis for a Fiscal Year 2001 funding request to be submitted by NOAA in 1999 with NASA support. Thus, the audience for this report is NASA / NOAA program management, Department of Commerce (DOC) staff, Office of Management and Budget (OMB) staff, and congressional staff. The report is also for use by NASA and NOAA in developing implementation approaches and cost estimates, as a resource for data center management, and by data managers and scientists involved with NASA and NOAA data management and user services activities.

1.3 Scope of the LTA Program

The scope of the LTA program comprises NASA EOS data and derived products and NEXRAD, GOES, NPOESS and other data sets / products needed to interpret and use EOS data/products. This includes any other satellite and *in situ* data sets needed for calibration, validation, and interpretation - especially between different instruments or members of a series of instruments flying on different platforms. The scope is recognized as an extension of current long-term archiving of satellite and *in situ* atmospheric and oceanographic data performed by the NOAA National Climatic, Oceanographic, and Geophysical Data Centers, (NCDC, NODC, NGDC respectively) and of cryospheric data by the National Snow and Ice Data Center (NSIDC). Similar archiving activity is conducted by the USGS's EROS Data Center (EDC) for land-surface data. EDC will assume future responsibility for NASA's EOS land-surface data.

The requirements for the LTA program developed by the workshop are stated in general terms applicable to all of the satellite and *in situ* data sets within scope of the program. The requirements are also applicable to the NASA/USGS land-surface data sets. These requirements constitute the first step toward a set of guidelines and principles for long-term archiving. They can be validated by comparison with successful long-term archives used currently and then used to gauge the performance of the LTA program for EOS and related data sets.

The LTA program is part of the nation's overall data management (collecting, processing, and disseminating operational and near-term research data and products, as well as storing, cataloging, preserving and providing access to Earth system data and information). Operation of the LTA program is embedded within a broader information system, with active interfaces to systems receiving data, products, and documentation and systems providing user access. Operational elements of the LTA program will extend beyond the data centers to include other institutions and individual research groups concerned with care for the data and service to the community. Provisions will be allowed for transfer of the data to a data center if an institution's or research group's interests change or for long-term preservation of the data. The LTA program fully supports the National Science Foundation's (NSF) standard policy for its environmental programs: that all data obtained in the course of agency-funded research is part of the public domain and is to be sent to the appropriate archives within a stipulated time period.

2.0 Guiding Principles for the Long-Term Archiving Program

Scientists at the workshop developed the following vision statement for the LTA program:

"A continuing program for preservation and responsive supply of reliable and comprehensive data, products and information on the Earth system for use in building new knowledge to guide public policy and business decisions."

The term "**continuing**" emphasizes that the LTA program must be essentially permanent. This does not mean that all data sets must be archived forever, but that the program itself must essentially continue forever.

Instead of a specific implementation plan, "**program**" suggests the creation of a more general document supported by a dedicated line item budget. This will allow someone to take official responsibility of the program, as well as to be accountable for its successes and failures. It is not, however, intended to imply a specific organizational structure; needs and goals can be met in a variety of places both inside or outside formal governmental structure – wherever people with the best capability, expertise and commitment reside.

"**Preservation**" emphasizes the critical importance of responsible care of the archive. Responsibility and authority for the LTA program will be vested in organizations that know about and care about the content of the archive - knowledgeable people with strong scientific interest in the data and all of its supporting information. These organizations will work to ensure the utmost support for present and future users of the archive.

The inclusion of "**responsive supply**" in the vision statement stresses the necessity for ready access to data, derived products, support documentation and user services in a manner that meets the LTA community's needs. One approach is to establish active mechanisms for securing continuing advice and guidance on the content and quality of the LTA's services.

"**Comprehensive**" is not meant to imply all long-term data sets should be preserved. Instead, it suggests inclusion of a carefully selected set of data, derived geophysical products, and

supporting documentation. Criteria for data selection will encompass scientific value, quality of the data and supporting documentation, relationship to other archived data sets, and resource constraints.

"**Building new knowledge to guide public policy and business decisions**" is an explicit recognition by the science community that the fundamental purpose of the LTA program is for the benefit its the investors, the American public. The goal is to incorporate improved understanding of climate change, environmental hazards and other issues to enhance public policy and assist in intelligent business decisions.

Implicit in the vision statement are two guiding principles:

The first principle is simplicity. At every step, the LTA program must be examined to assure clarity and comprehensibility while still meeting program goals and user needs. Activities ranging from review of archive contents to details such as data file structures and related access software should be evaluated.

The second principle is endurance. The LTA is a resource for future generations of scientists and citizens who are not here to speak for themselves; their needs must be anticipated. Steps such as the design of comprehensive and complete documentation, investment in calibration, validation, science assessment, and data quality all need to be of concern. Consider questions a future user, scientifically knowledgeable but unfamiliar with the data set, would be faced with:

- Is the LTA be readable? (e.g. was the data set written on a standard medium readable by standard equipment, or was it stored on custom equipment only available through a single vender? Was it backed up securely?)

- Is the documentation sufficient, especially given that the people originally involved with the data set are either be gone or no longer familiar with it?

- Are the data or product formats and structures clear?

- Does the archive user support staff have in their possession the information they need to answer user questions about the data set or access to it?

3.0 Intent of the Long-Term Archiving Program

Three issues are addressed in this section. In response to the LTA program vision statement "to guide public policy and business decisions", the first section discusses the broad public interests served by the LTA program. The second section discusses immediate scientific applications to support the "building new knowledge" part of the vision statement. These sections are followed by a segment on the scientific consequences should the LTA program not be implemented. Finally, the close relationship between the LTA program and the needs of the U.S. Global Change Research Program are discussed.

3.1 LTA Program for Public Benefit

The ultimate purpose of the LTA program is to contribute to the protection and health of the national economy and the enhancement of the quality of life of the American public. The intent is to ensure the American public obtains the greatest possible return on the enormous investment into the acquisition of Earth science data – accompanying the already valuable operational and short-term data usage.

Public benefit is accomplished through the LTA program's use by the scientific community and others, for it is scientific research, where archived data and information is essential, that yields improved understanding of the Earth system and global climate change. An increased awareness of our environment can then be used as the basis for different economic actions such as assessments of global or regional socioeconomic impacts of climate change, assessments of environmental hazards and their impacts, and guidance for formulation of public policy by federal, state, and local governments from the assessments' conclusions. At the same time, application of such research on a regional and local scale can provide valuable input into investment decisions by businesses and individuals in areas such as real estate, agricultural planning, recreational facility planning, energy use projection, etc. The LTA program can establish a reliable, long-term record of the state of the Earth system that can be translated by research and analysis into guidance for more effective management of national, business, and individual resources.

3.2 Scientific Objectives

As stated earlier, the basic scientific purpose of the LTA program is to enable and facilitate the best possible science and highest-quality assessments as this work will be used for making policy and business decisions. To "enable and facilitate" means to do more than make such research physically possible – it includes taking steps to make research easier or more effective by optimizing program services for the convenience of the users, primarily by making those services simple, straightforward, direct, and responsive.

A second scientific purpose of the LTA program is to document Earth system variability and change on global, regional, and local scales, building and maintaining a high quality base of data and information and establishing the best possible historical perspective critical to effective analysis and prediction. Comprehensive understanding of global climate change requires the longest possible time series of the highest quality. Subtle features in climate trends and variability can often only be detected in the longest and best geophysical records. Thus, proper preservation of these records are critical in the determination of any affects anthropogenic activity may have on the Earth system. The value of the LTA data and information base grows as it is extended in time and enhanced in quality. Also increasing its value is inclusion of datasets developed by process studies such as TOGA-COARE (in addition to the products of routine operational satellite and *in situ* data collection programs).

A key scientific objective of the LTA program must be to ensure archive holdings, facilities, and services are actively promoted and made readily available to the maximum number of users. The

best possible scientific understanding of any issue can only be reached when all opinions and ideas are explored. Therefore inaccessibility, complexity and cost should not be obstacles in the use of the LTA.

A fourth scientific purpose of the LTA program is to enable and facilitate future research. The Earth system data and information base must be carefully preserved, selectively extended, pruned, and maintained over time. The archive will be a valuable resource for problems not yet recognized. Thus, the LTA program is an investment whose return will also be seen in the future. Examples of future uses include:

- The emergence of new problems in public policy or business planning that will have policy makers and decision makers looking to the science community for new guidance.

- The re-evaluation of current theories as continued research reveals initial misinterpretation. This may require looking at existing data in new ways, and/or drive the acquisition of entirely new data that must be correlated or compared with existing data.

- The development of new scientific theories leading to entirely new geophysical products or improved versions of existing geophysical products and/or development of new geophysical products by combining old and new data or by using new combinations of concurrent data from multiple sources (e.g. improved four dimensional data assimilation).

- The development of new approaches and capabilities for scientific computation that enable and stimulate currently impossible ways of working with high volume, long-term data sets. The very recent explosion of inexpensive but powerful workstations and personal computing technology along with the emergence of the internet have vastly expanded data access, data handling, and computational capabilities available to the average research user. Large technological advances are expected to continue, making enhanced capabilities a feasible reality.

3.3 Consequences of Not Having the LTA Program

Without an LTA program in general, the objectives just discussed could not be fulfilled. Many scientific questions relevant to key public policy issues and business decisions can not be addressed without making use of long time series. Without a comprehensive, reliable and accessible data archive, the majority of scientific studies would be limited to research focused on the available short-term data. This is an extreme situation. In reality, many data sets are currently being archived, preserved, maintained, and made available to researchers and other users. However, without the extended LTA program discussed at the workshop and in this report, there would be:

- No long-term archiving of the forthcoming NASA EOS data and other related NASA Earth science data. While NASA's Distributed Active Archive Centers (DAACs) have the majority of, if not all of the required capabilities, NASA has neither the mission nor the budget for long-term archiving of its Earth science data.

- No long-term archiving of NPOESS data. NOAA has the mission to archive the data, but no resources in its budget.

- Only a very minimal capture of low-level data from the GOES satellite series and the NEXRAD (WSR-88D) Doppler radar network. NOAA again has the mission to archive the data, but resources that only cover bare capture and limited access to these two large data sets

- Severely compromised long-term value of key data set collections developed by organized process studies or field campaigns (e.g. TOGA-COARE). The LTA program is needed if the preservation, documentation, and long-term availability of these collections are to be continued.

- No coordination and management of long-term archiving as a program, no focal point for responsibility and accountability, and no formal way to engage the participation of the science community.

Lack of a formal LTA program would affect many important projects such as the study of temporal dynamics and the evolution of climate change, leading to poor scientific contribution to policy issues and business decisions important for the future health of the economy.

3.4 Relation to U.S. Global Change Research Program (USGCRP)

A key programmatic requirement for the LTA program will be to provide support to the USGCRP. The central objectives of the USGCRP are:

- To observe and document changes in the Earth system;

- To understand why these changes are occurring;

- To improve predictions of future global change;

- To analyze the environmental, socioeconomic, and health consequences of global change;

- To support the foremost assessments of global environmental change issues.

The USGCRP supports a number of integrative, multi-agency efforts that contribute to its scientific goals. The overarching objectives include:

- Ensuring a long-term, high-quality record of Earth's dynamic system including natural variability and change

- Providing all users ready and affordable access to the full spectrum of global change data, products, and information in a useful form.

- Gaining an understanding of the interactions among the physical, chemical, geological, and biological processes that define Earth's trends and fluctuations on global and regional scales.

The LTA program directly prescribes to the first two objectives of the USGCRP with the intent of accomplishing the third.

In addition to the USGCRP and LTA program's concordant purpose, the LTA program also supports many of NASA's own research goals, the data management dimension of the IPCC's requirements for climate monitoring, and the recommendations of the Workshop on Long-Term Climate Monitoring (see Appendix C).

4.0 Lessons Learned from Past Experiences with Data Archives

The importance and relevance of a long-term archiving program can be demonstrated best through past experiences. These examples, which illustrate the points discussed in the previous section, provide both positive and negative lessons emphasizing the importance of a well-planned LTA program. They are grouped into four sub-sections:

4.1 Examples of scientific success due to LTA programs: Cases where the existence of a data archive has allowed reprocessing to produce new products for global change research.

4.2 Examples of scientific success due to LTA programs: Cases where the existence of a data archive has allowed pursuit of previously unanticipated applications.

4.3 Examples of scientific limitation and failure due to poorly structured LTA programs: Cases where the lack of fully comprehensive data archives have severely limited their use for scientific research.

4.4 Examples of new scientific opportunities made possible by an LTA program: Scientific questions and hypotheses whose solutions require LTA program services.

4.1 Examples of scientific success due to LTA programs: Data Reprocessing

Each of the following examples illustrates a case where a significant scientific accomplishment was only possible because of an existing long-term archive of one or more necessary data sets.

They demonstrate that archiving Level 1 data and including complete documentation will make the successful generation of new geophysical products possible. The examples also prove the advantages of storing data in a form that facilitates multiple reprocessing, an inevitable action as new product algorithms are developed and old ones refined.

4.1.1 Ozone Trends and Tropospheric Aerosols

Ozone is an important trace gas. Not only is the climate system sensitive to changes in the vertical distribution of ozone in the troposphere and stratosphere, but decreases in stratospheric ozone allow greater penetration of solar UV emission to the Earth's surface, potentially impacting organisms living there. Thus, ozone profile information, particularly long-term trends, is an important aspect of climate change modeling studies (*Karl et al.*, 1995).

Total ozone has been measured for more than thirty years by a network of about 200 ground-based Dobson spectrophotometers, reliant on one particular instrument maintained by NOAA as a calibration benchmark. In November 1978, NASA launched the Nimbus-7 satellite with the first Total Ozone Mapping Spectrometer (TOMS) instrument. TOMS operated for almost fifteen years (until May 1983), providing daily maps of ozone for most of the globe. A second TOMS was flown aboard the Meteor-3 satellite and operated from August 1991 until December 1994, and a third TOMS was flown aboard the ADEOS-1 satellite operating from August 1996 until June 1997. A fourth TOMS was launched aboard an Earth Probe mission in July 1996 and remains operational.

During analyzation of the TOMS data, values of total ozone less than 180 Dobson units were routinely flagged by the processing software as physically implausible. In 1984, however, new, reliable ground-based measurements in Antarctica showed ozone values could in fact fall to this level (*Farman et al.*, 1985). The question now faced NASA scientists: were the low TOMS measurements valid or were they the result of an instrument or calibration problem?

NASA investigated the situation and concluded that the previously disregarded measurements were real. A report followed showing the re-calculated TOMS measurements indicated ozone depletion over almost all of Antarctica (*Stolarski et al.*, 1986). After re-evaluating past data, the ozone deficiency was found to be detectable as far back as 1976. Striking images of the TOMS ozone analysis were produced and the term "ozone hole" entered the popular vocabulary. Since then, repeated reprocessing of the TOMS data has allowed significant improvement in our knowledge of stratospheric ozone levels. Figure 1 shows the ozone hole's progressive development as discovered throughout the TOMS data record.

Since this discovery, ozone depletion over both hemispheres has been actively monitored. The data now benefits from the intercomparison of several TOMS instruments while the gap in TOMS coverage (between December 1994 and July 1996) is filled by SBUV/2 data and European Global Ozone Monitoring Experiment (GOME) data. An aggressive effort has also been made to compare the TOMS ozone measurements with ground based measurements (e.g. *McPeters and LaBow*, 1996) and to derive total ozone maps from earlier satellite data (e.g. *Stolarski et al.*, 1997). Results of the improved processing algorithm to Dobson ozone data can

be seen in Figure 2, and Figure 3 shows the comparison between ozone measurements from the SBUV/2 instrument on NOAA 11 and Dobson ground based data.

The value of data reprocessing has also been demonstrated with other ozone and atmospheric chemistry data sets. Archived seventeen-year-old SAGE-I (Stratospheric Aerosol and Gas Experiment) data, for example, is currently being reprocessed to enable intercomparison with more recent SAGE-II data. The SAGE-II data has also been reprocessed a number of times, including a re-derivation of ozone profiles from Level 1 data (*Rusch et al.*, 1998) which showed improved agreement with balloon ozonesonde profiles in the lower atmosphere. SAGE-II data is currently being reprocessed again to improve the algorithm used to correct for the effects of high atmospheric concentrations of aerosols due to discreet events (such as volcanic eruptions). In other cases, Nimbus-7 SBUV and UARS HALOE (Halogen Occultation Experiment) data have been reprocessed at least five times, ATMOS (Atmosphere Trace Model Spectrometer) data is on its second reprocessing, and as soon as a major revision to the data-processing algorithm is completed, twenty-year-old Nimbus-7 LIMS (Limb Infrared Monitor of the Stratosphere) data will also be re-evaluated.

The continuing ozone and aerosol efforts demonstrate several important lessons. First, preservation of low-level data and the associated documentation allows for continual re-processing as scientific understanding and abilities improve. Also, intercomparison and intercalibration of data from multiple sources is only possible with this kind of preservation; the highest quality products rely on data from multiple sources to help mitigate the effects of instrument biases and to provide the most complete coverage possible. A third lesson is the ability to use the data to retrospectively generate additional, unanticipated scientific products. Though not mentioned explicitly in this section, measurements such as the TOMS aerosol index, surface UV flux and reflectivity, and sulfur dioxide can now be calculated using the archived data.

4.1.2 Precipitation Trend Analysis

Detecting trends in precipitation is another way to study climate change and its environmental effects. A variety of different instruments and observation techniques have been used over the past hundred years or so. Modern measurements now include data from weather satellites. For both *in situ* and satellite data, information about the instruments and how they were used is essential for proper interpretation and analyses of the precipitation data they collected. For example, Figures 4 and 5 are 98-year-long precipitation trends based on *in situ* data published in the 1995 IPCC report. Figure 6 shows the information on how the *in situ* observations were made, the changes made in observation techniques and instrumentation over time, and the variations in the precipitation measurements that resulted. Because information on the different instruments and observation techniques was available, scientists could account for disparities in methodology and produce a reliable record of precipitation trends.

Satellites used for detecting precipitation data (or any other geophysical parameter) are not immune to these problems either. One current problem is the gradual drift of the time of day that 'sun synchronous' polar orbiting satellites actually view an area. The top panel of Figure 7 shows the time-drift in the daytime equator crossing for NOAA's series of polar orbiting

satellites. Changes in the time of day of the satellite overpass can produce false trends that look like climate change, but are in fact just artifacts of the satellites' orbital characteristics. It is as if a series of daily thermometer readings are made, first at 8:00 am but slightly later each day until the last one is made ten years later at 12:00 noon. Analysis of the series would reveal, amidst seasonal cycles and the ups and downs of passing weather, an underlying trend towards warmer temperatures not actually due to a change in climate, but due to the changing time of observation. The false trend could also bury any signals of true climate changes that might have occurred. In the case of satellite-measured long wave radiation emission data, changes in the radiation values track the orbital drift of the satellites (Figure 7). Since this data is often used for detecting precipitation intensity, this drift, if not corrected, would effect calculations on precipitation intensity, and lead to false trends.

This example demonstrates the importance of preserving all available documentation on instrument characteristics, observation methods, etc. for both *in situ* and satellite data. This is the only way true, climate trends can be confidently separated from artificial ones.

4.1.3 Global Near-Surface Temperature Analysis

Global near-surface temperature trends are potentially clear indicators of climate change, provided that temperature trends due to climate can be distinguished from trends due to other causes. Discrepancies between temperature trends determined from weather station data at the Earth's surface and atmospheric temperature trends determined from satellite microwave instruments has led to efforts to refine both data sets. Urban heat islands were identified as one possible influence on land-surface weather station measurements. The weather stations were separated into two categories, stations in the vicinity of large urban areas and those in rural regions, and visible light measurements from DMSP OLS satellite data (section 4.2.2) was used to examine the discrepancy between ground and satellite surface-temperature trends (Figures 8 and 9).

In this example, archives of data from both satellites and land-based observations were evaluated using another data set in order to confirm their accuracy and to investigate an unanticipated discrepancy. The archived data and documentation proved a valuable resource because it was stored in a manner that facilitated the comparison of multiple records in a way not originally intended.

4.1.4 Ocean Frontal Analysis

Ocean fronts are relatively narrow zones of enhanced horizontal gradients of physical, chemical and biological properties. As a consequence they are also regions in the ocean that separate broader areas of differing vertical structure. These frontal regions are important both because of the role they play in ocean dynamics and because of the assistance they provide in characterizing water mass boundaries. Their dynamical importance stems from their association with strong currents, such as the Kuroshio or the California Current, and their association with regions of large vertical motion such as that observed in the subtropical convergence zone. Additionally, fronts are regions of large horizontal and vertical mixing.

The large vertical mixing results in an increase in the concentration of plankton at fronts. As a consequence, fronts play a significant role in the world's fisheries; the richest fishing regions in the world ocean are also regions of significant frontal activity. Until recently, however, a multi-year time series with which to investigate the relationship between ocean fronts and the worlds' fisheries has not existed. Nor has the data existed to provide a robust estimate of the impact of changes in the global climate on the density of fronts in the global ocean. This situation is changing. NOAA has maintained a long-term archive of AVHRR data from which SST (sea surface temperature) fields needed to detect ocean fronts can be derived. NOAA and NASA sponsored a project known as the NOAA/NASA AVHRR SST Pathfinder Project. The goal of the project was to produce a consistent global SST time series from five channel AVHRR data, spanning June 1981 to the present. Processing is completed for data from 1985 to 1996, yielding daytime and night time SST fields at 9.28 km resolution. The SST algorithm is tuned with *in situ* measurements. A multi-image edge (SST front) detection algorithm has been applied to the series of processed images providing for the first time global maps of the probability distribution of fronts in the world ocean. The algorithm divides each image in 32x32 pixel squares and performs a histogram analysis of each square looking for multimodal distributions. Figures 10a and 10b are examples of two of the maps obtained by combining the output from this algorithm applied to the 11 year time series. They are for the South Atlantic, one for January and the other for July, and they show quite clearly the differences that occur in frontal density from austral summer to austral winter.

The Pathfinder Project was possible due to the existence of archived Level 1 AVHRR data. The success of the project underscores the importance of retaining Level 1 data in a long-term archive and making reprocessing significant volumes of data feasible.

4.1.5 Snow Cover Analysis

Seasonal changes and trends in snow cover extent provide an important source of climate change information. Snow cover extent, measured by satellite data over the last 20 years, appeared to be decreasing during the spring and summer months. NOAA produced snow cover extent information from VHRR (Very High Resolution Radiometer) and AVHRR (Advanced VHRR) satellite instruments from 1972 to the present. As the instruments changed and understanding of the underlying science improved, different algorithms were used to filter the data. This resulted in an inconsistent record where any natural climate changes were masked by false shifts due to the different algorithms (particularly changes in a land-ocean mask and an undocumented change in post-processing of the snow cover product). The Intergovernmental Panel on Climate Change (IPCC), in their 1990 report, used the miscalculated data to erroneously report a very large reduction in snow cover. Fortunately, the data from which information was derived had been preserved in a long-term archive. NOAA was able to reprocess the data using an improved and consistent algorithm over the entire record. The result was a significantly refined long-term record of snow cover extent (Figure 11) and the IPCC published the corrected data (Figure 12).

The importance of a long-term archive is deftly illustrated in this example. Climate change is a subject of global concern and debate. Statements made by a panel as distinguished and prominent as the IPCC are far-reaching, and must be as factual as scientifically possible. Because the VHRR and AVHRR data was properly archived, the IPCC was able to revise its

publication and retain credibility. This example also highlights another point already made in previous examples: a long-term archiving program must contain a diverse set of data. Only by comparison is it possible to detect artifacts due to incorrect analyses.

4.1.6 Ocean Topography

Data from radar altimeters flown aboard Seasat, Geosat, ERS-1 and later satellites have been used to map ocean surface topography and derive other parameters potentially useful for detecting the effects of climate change. The ability to distinguish such small changes over a long period of time requires an accurate, long-term data set. The greatest source of error in these calculations were satellite position, a parameter critically dependant on knowledge of the Earth's gravity field. As better information on the gravity field became available, data from the earlier Seasat, Geosat, and ERS-1 satellites were reprocessed yielding higher quality geophysical results. The older, reprocessed information can now be used with ocean surface topography information produced from more modern and sophisticated satellites to generate a longer, more accurate data set.

Again, this example emphasizes the importance of preserving both the data and supporting information in such a way that it can be easily re-evaluated. This is particularly crucial for the creation of very long time series, where constantly changing technology means that periodic re-evaluation of the data is almost certain.

4.2 Examples of scientific success due to LTA programs: Unanticipated Applications.

Each of the following examples presents a case in which completely unexpected but scientifically significant uses were made of a data set secured in a long-term archive. The beneficial results from unanticipated data manipulation underscores the scientific flexibility that preservation of Level 1 data and complete documentation provides.

4.2.1 NOAA AVHRR

The NOAA AVHRR instrument was designed to produce cloud imagery and sea surface temperature fields to support operational weather forecasting by the National Weather Service. AVHRR data is now applied to many other unanticipated research areas, such as vegetation dynamics (e.g. the Normalized Difference Vegetation Index (NDVI) product computed from the uncalibrated AVHRR visible channels), drought and vegetation stress, carbon dioxide uptake, carbon budget parameters, and fire susceptibility / risk in the U. S. Great Plains and elsewhere. AVHRR data is also used to study ice mass balance - to detect ice movement, calving of large ice bergs, and to measure ice surface temperature.

A complete archive of global Level 1 AVHRR nominal 4 km data (GAC - Global Area Coverage) was established and has been maintained by NOAA ever since the first AVHRR instrument flew on the TIROS-N satellite. The first five years of AVHRR GAC Level 1 data (along with AVHRR 1 km resolution data sets and TOVS Level 1 data sets) were collected on high density tapes by a system which was also (and primarily) used as an on-line buffer between

the real-time preprocessors that generated the Level 1 data from raw data and the mainframe computers that produced geophysical products from the Level 1 data (particularly sea surface temperature fields from AVHRR data and vertical profiles of atmospheric products from TOVS data). NOAA decided to upgrade the operational processing system and eliminate the need for the high density tape system as an on-line buffer. A time window, originally sufficient but which eventually proved to short, was allowed to migrate data from high density tapes to 6250 bpi tapes. Once the flow of incoming data was switched to the upgraded system, NOAA began to migrate the archived data. It became apparent that the effort would only partially succeed. Luckily, to prevent losing the rest of the data, NASA was able to provide funding for NCAR to finish the migration. When the NOAA system deactivated, they were able to send the remaining, untransfered tapes to NCAR and the archives were salvaged. The migration problem would not have occurred had there been a program focus for data management within NOAA to advocate the importance of retaining the AVHRR archive and to secure the resources and time needed to accomplish an orderly migration of data.

Over the years since the AVHRR instrument was first flown, a number of data reprocessing activities have been performed. Notable among these have been the NOAA / NASA cooperative AVHRR Pathfinder projects, which reprocessed more than 14 years of AVHRR Level 1 data. Consistent sets of ocean, land and atmospheric data are now available for use in global change research. Groups of scientists guided the Pathfinder efforts, including approval of the algorithms used to generate the geophysical products. The first step, to obtain the Level 1 data and documentation and transfer it to a medium suitable for reprocessing, took significant effort.

The reprocessed AVHRR data is being used in a number of ways. The ocean data is being used for ocean-front analysis (4.3.2), and an AVHRR Land Pathfinder project produced daily and ten-day composite reflectance and brightness temperatures along with NDVI, cloud identifiers and ancillary data. NOAA is conducting the AVHRR Pathfinder Atmosphere (PATMOS) project to reprocess the AVHRR Level 1 data to correct radiances for calibration drift. The corrected records include cloudy and clear radiance statistics, total cloud count, top of the atmosphere radiation budget (cloudy and clear sky) and aerosol optical thickness over the oceans. These daily, five-day, and monthly averaged products will be used in climate change studies. Last, the NOAA / NASA AVHRR Polar Pathfinder project is reprocessing the Level 1 AVHRR data set to produce twice-daily ice surface temperature and albedo measurements and daily ice velocities for the polar regions.

Several lessons can be learned from the AVHRR program. The lack of a program specifically focused on long-term data management nearly led to a disaster - loss of a priceless archive of scientifically significant global change information. The continuing success of the AVHRR Pathfinder projects demonstrate the often unforeseen value of archived data. Only with such an LTA program can a long and consistent data set be continually reprocessed for use in unanticipated and exciting ways. The importance of documentation is underscored in the NOAA Polar Orbiter Data Users' Guide (*Kidwell*, 1997, the most recent edition).

4.2.2 Defense Meteorological Satellite Program - Operational Linescan System

The Operational Linescan System (OLS) instrument package has been flown aboard the Defense Meteorological Satellite Program (DMSP) Block 5D-1 and 5D-2 series of polar orbiting weather

satellites since 1976. OLS includes two telescopes and a photo-multiplier tube. Visible and infrared (IR) imagery from DMSP OLS instruments are used to monitor the global distribution of clouds and cloud top temperatures twice each day, during the day and at night. The archive data set consists of imagery from low resolution (2.7km) global coverage and high resolution (0.55 km) regional coverage. Night-time visible data is obtained from the photo-multiplier tube, designed to produce cloud imagery by moonlight

NGDC has archived analog images of the aurora from DMSP OLS sensors since 1972. The National Snow and Ice Data Center (NSIDC) has maintained the archives of other analog OLS data since 1979. In March 1992, NGDC began to receive the full DMSP digital data stream (including OLS, SSM/I, and other instrument data).

The archived OLS data are now used for a variety of other studies in addition to providing cloud imagery and cloud top temperatures, its original application. Ground-based light sources such as fires, natural gas flaring and urban lights are measured for information on natural fires, biomass burning, power failures and population density (4.3.1). OLS data are also used to survey atmospheric light sources such as lightning and aurora, to monitor fisheries operations and to examine global vegetation and sea ice distribution and movement.

The OLS example again emphasizes the many unexpected benefits of a well-maintained data archive. It's easily accessible digital format boosts its value even more.

4.2.3 NOAA TOVS - Upper Troposphere Humidity

Level 1 data from the High Resolution Infra-Red Sounder (HIRS), part of the TIROS Operational Vertical Sounder (TOVS) package flown aboard NOAA operational polar orbiting satellites, has been archived since late 1978 when the first HIRS was flown aboard TIROS-N. The algorithm for deriving upper tropospheric humidity (UTH) from HIRS data, however, was not developed until 1996, eighteen years later. Data from the TOVS water vapor channels had not been extensively used for several reasons: 1. the quality of *in situ* data from radiosondes, used for validation of the TOVS measurements, was poor, 2. the inability of early atmospheric models to properly incorporate water substance vapor, and 3. the lack of knowledge of the spectroscopic details of water vapor radiative transfer. Improvements in each of these areas and new interest in the role of water vapor in global warming led to a re-examination of the TOVS water vapor channels in the early 1990s. The re-examination was highly successful and resulted in breakthroughs for both numerical weather prediction and climate analysis. The potential significance a global history of UTH has for climate change research brought on a major effort to develop consistent inter-satellite calibration for the HIRS instruments flown on TIROS-N, NOAA-6 and succeeding NOAA series satellites. The culmination of this effort was an analysis of the annual UTH cycle spanning a time period of 1979 through 1995 (Figure 13, *Bates et al.*, 1996).

4.2.4 NOAA TOVS Radiances

The original concept for TOVS data was to derive vertical profiles of atmospheric temperature and humidity to mimic the profiles produced by balloon borne sensors and for use in research and operational numerical forecasting models. Many years after the first TOVS package was

flown, scientists determined the original approach could be improved through techniques developed from work with calibrated TOVS radiance data (e.g. *Eyre et al.*, 1993). A long-term archive of TOVS Level 1 data had been preserved, beginning with the first data collected in late 1978, allowing scientists to go back and develop a new time series of TOVS instrument radiances from the original Level 1 data. The improved results greatly enhanced the TOVS value for long-term studies and climate models; they have been used in the European Center for Medium-range Weather Forecasts (ECMWF) reanalysis from 1979 to 1993 and are planned for use in other reanalyses of global atmospheric observations.

4.2.5 DMSP SSM/I and NASA SMMR

The Special Sensor Microwave / Imager (SSM/I) instrument, flown aboard DMSP polar orbiting weather satellites since June 1987, was originally intended to support operational weather forecasting by the U.S. Air Force and Navy. At that time there was a limited effort to develop research algorithms to monitor surface processes. SSM/I data are now used for a wide variety of global land, atmospheric and ocean surface applications. The SSM/I is a well-calibrated passive microwave radiometer that is particularly suited for measuring hydrologic cycle parameters under all weather conditions. These include measurements of snow cover, frozen ground, land classification, sea ice, wind speed over the oceans, cloud liquid water and water vapor, and rainfall. Figure 14 is an example of an SSM/I water vapor product.

The SSM/I instrument series (six instruments have flown, with a seventh scheduled for future flight) will continue at least through the year 2000 before being replaced by its successor, SSM/IS, which is a combined passive microwave imager and sounder. The SSM/I is itself a successor to the earlier Nimbus-7 Scanning Multichannel Microwave Radiometer (SMMR), which collected data from 1978 through 1987. The combined SSM/I - SMMR time series exceeds 20 years and provides an invaluable, albeit unanticipated, all-weather data set for use in global change studies.

Because of their importance for global change research, the SSM/I and SMMR data sets were made the object of a NOAA-NASA Pathfinder Project. The project's goal was to produce a carefully calibrated, consistently processed, thoroughly documented long-term time series (20 years and growing) data set of SMMR and SSM/I brightness temperatures. The new SMMR-SSM/I brightness temperature data set is now available to the science community for present and future use as the basis for a variety of geophysical products. Figure 15 is an example of one such product generated by NSIDC, a 20-year time series of depicting the decreasing trend of global snow cover extent.

Production of the Pathfinder data set, and the consequent derivation of geophysical products, was only possible because the low-level SMMR and SSM/I data sets, calibration information and other documentation had been well preserved.

4.3 Examples of scientific limitation and failure due to poorly structured LTA programs: Incomplete Documentation

When considering the requirements and essential functions of a LTA program, particular attention must be paid to long-term archiving efforts that have encountered problems that seriously undermined the usefulness of the archived data to the research user community. The three cases are discussed in this section each began as the best approach possible for the resources available at the time. Each has encountered problems that the LTA program must avoid.

4.3.1 NEXRAD (WSR-88D)

The WSR-88D (Weather Surveillance Radar - 1988 Doppler) network, established by the Next Generation Weather Radar (NEXRAD) program, is a joint effort of the Departments of Commerce (DOC), Defense (DOD) and Transportation (DOT). Network operations are managed by the National Weather Service (NWS), U.S. Air Force (USAF) and the Federal Aviation Administration (FAA), respectively. The WSR-88D network includes 158 radar data acquisition stations positioned across the United States, including Alaska and Hawaii, and at a few international sites. The WSR-88D radar network generates a large volume of base data (over 100 terabytes annually), comprised of reflectivity, velocity and spectrum width (also referred to as Level II data). Up to 80 products, collectively referred to as Level III data, can currently be derived from the Level II data and are used to aid operators in monitoring weather conditions and preparing forecasts and advisories. The prime mission of the radar system is to support real-time operational severe weather analysis, forecasting, and warnings. The focus of the NEXRAD program is almost exclusively on the highly critical real-time use of the data and their products. The original plan called for archiving only 22 of the 80 Level III products at only 120 NWS sites and there were no plans to archive the low-level (Level II) data from any site. Shortly before network implementation, it became apparent that the Level II data would be invaluable not only for development of new or improved operational products but also for long-term climate change studies. Agreement was reached among the three agencies to archive the Level II data at all 158 sites at the discretion of the station manager, though unfortunately, a well-planned archive was not discussed.

At the present time, with all the 158 WSR-88D sites in operation, Level II data are transferred to the NOAA National Climatic Data Center (NCDC) in Asheville, North Carolina on 8mm tapes, at a rate of more than 2,000 tapes per month. The tapes are cataloged and placed in environmentally controlled storage. Although the data is being captured, the slow transfer rate (about 1GB per hour) and fragility of the 8mm tape drives causes user access to be painfully slow and expensive. Copying one tape with 4.5 GB of Level II data - a period of about two days - requires 4 to 5 hours if no transfer problems occur. A user request for a substantial volume of the data is very expensive and data transfer is extremely time-consuming. These factors combine to make the access operation very labor intensive and expensive. From the user's perspective, the cost trade-off between storage and access was well planned, making access to the data very difficult. In addition, a large volume of data is rapidly being accumulated and migration to an archive medium both more secure and offering responsive user access will be an expensive, time

consuming effort. The migration effort can not be avoided, however, if the data is to be preserved and its value realized.

The primary lessons learned from the NEXRAD experience are:

1. For full advantage to be taken of an archive, life-cycle data management is needed, beginning at the planning and analysis stage and continuing through development and implementation. It should involve researchers and data managers as well as data collectors and primary users.

2. Data managers have to be more aggressive, especially in the early stages of development. Alternative uses of data must be fully explored from a researcher's perspective as well as the immediate operational user or principal investigator / research team.

3. Archive media must be carefully chosen, with thoughts on future migration as well as current and projected user access requirements.

4.3.2 NOAA GOES

The series of geostationary meteorological satellites were launched in 1966 as the NASA Applications Technology Satellites, continued with two Synchronous Meteorological Satellites and were followed by the operational NOAA Geostationary Operational Environmental Satellite (GOES) series. The concept and key instrument on GOES, the Visible and Infra-Red Spin Scan Radiometer (VISSR), was originally conceived and developed for NASA by Verner Soumi and his colleagues at the University of Wisconsin's Space Science and Engineering Center (SSEC). By 1979, SSEC had also developed a means of recording (at the time) a very high volume of GOES digital data to preserve for research use.

In 1977 NOAA began to plan operational long-term archive capability for GOES data and won approval for a FY80 budget initiative to fund the plan. After examination of commercially available alternatives, NOAA turned to the SSEC and their system as the only affordable alternative and contracted them to construct the NOAA long-term archive for GOES data.

At present, SSEC holds 50,000 cassettes containing Level 1 GOES visible and infrared data for the period 1979 through September 1997. The nature of their archive system is similar to the NEXRAD archive discussed in 4.3.1 - retrieval of substantial quantities of GOES data for climate research is a painfully long and prohibitively expensive task due to the tape system used by SSEC. As a result, the use of archived GOES data has been almost exclusively limited to case studies of narrow scope such as severe weather phenomena. The scientific value of full resolution GOES data for climate research is largely unknown, and the potential value for information in areas such as wind - sea surface temperature relationships or hurricane systems is untapped. The only systematic effort to use GOES data for climate studies has been through the International Satellite Cloud Climatology Project (ISCCP) which collects spatially and temporally sampled GOES data and related data sets to produces global cloud cover climatology products.

The GOES archive has fallen into a vicious cycle: because the GOES data is difficult and expensive to access and use, little research, virtually none of it on climate, has made use of the archive. As a result, the science value of the data is not proven making it impossible to win funding for migration to a more accessible system. In addition, while the present archive medium is very durable, the oldest cassettes are approaching the end of their useful lifetime. If the data are not migrated soon they may eventually be lost.

Lessons learned from the GOES archive experience include:

1. Ready and responsive access to archived data is a key factor in securing its preservation. Only through use is the scientific value of data proven and the expense of preserving it justified. The LTA program must regard promoting use of the data as a part of its data preservation strategy.

2. Life-cycle data management must be an ongoing process - it can not stop when initial implementation is made, even if the initial implementation is the best that can be done at the time. Planning for and accomplishing migration to a system offering improved user access is as important as maintaining the archive.

4.3.3 Seasat

Seasat data provides an example of the danger in an obsolete and unsupported media type. Seasat data was migrated by JPL's NASA Ocean Data System (NODS) from magnetic tape to optical disk in 1989. At the time, these optical disks provided a good, cost effective media renewal solution. Unfortunately, this media proved to be an inadequate long-term solution for two reasons. First, there was (is) no vendor-independent format for optical disks. Consequently, NODS became tied to a particular optical disk equipment vendor. Second, the software for the optical disks were written for machines that advanced technology has since made obsolete. Today the JPL DAAC, the successor to NODS, must maintain the obsolete system for the sole purpose of being able to read the disks. The DAAC cannot even upgrade the system because of the incompatibility with the optical disk driver software. So, while the decision to migrate the data to optical disk saved resources in the short run, the DAAC must now migrate the data again for it to be of any use.

Lessons learned from this experience include:

1. The LTA program must strive to maximize the use of standard, vendor-independent technology whenever possible, especially for data storage and any other elements of a system where stability is desirable and reasonable.

2. The use of a non-standard or vendor-dependent approach should be accepted only if a net advantage over a designated time period is clear, and after a thorough evaluation of other options has been made.

3. Costs for migration should be put into the program's budget from the outset

4.4 Examples of new scientific opportunities made possible by an LTA program: Future Scientific Research

Each of the following examples points to a need for long-term archiving as well as the benefit of combining data sets. They emphasize the importance for preserving the full documentation necessary to support intercalibrations and intercomparisons of various data sets.

4.4.1 Is the Earth's Troposphere Warming?

Answering this question requires a set of measurements made over a sufficiently long period of time. One such data set (*Christy*; 1995) is tropospheric temperatures derived from upwelling microwave radiation. These measurements were made by the Microwave Sounding Unit (MSU) flown on the NOAA series of polar orbiting satellites beginning in late 1978. The MSU measures radiances near 60GHz, a frequency range in which oxygen molecules radiate as blackbodies. Since oxygen is well mixed in the troposphere, its measured radiation provides an overall indicator of the tropospheric temperature. The algorithms for computing temperatures, contamination corrections, instrument calibration and orbit decay effects have been gradually refined to produce reliable records of tropospheric temperature (Figure 16). Over the past several years, as the algorithms have been improved, the MSU data has been reprocessed, and significant changes in the calculated estimates of global tropospheric temperature change have resulted.

This ongoing research has only been possible because the Level 1 MSU data and all of the information needed for its interpretation have been collected, archived, and made readily accessible.

4.4.2 Is The Biosphere's Productivity Declining?

The health of the biosphere depends on interactions with the hydrosphere and atmosphere. Changes in either of these will affect the productivity of the entire system, possibly in a detrimental way. Early warning of harmful modifications is fundamental to ensuring stability of global food supply and ultimately, the feasibility of continued life on Earth. It is therefore essential to monitor the biosphere with regard to such global change issues as climate, soil quality, soil chemistry, land use cover, precipitation, snow cover extent, biodiversity, land resource management practices, and atmospheric chemistry - particularly concentrations and distribution of gases that affect the biosphere. Is the climate changing in such a way as to cause measurable impacts on productivity? Is it affected by altered land cover (for example the geographical extent of different land cover types)? How about different atmospheric chemistry?

Such questions can only be addressed effectively using global data sets extending over long periods of time. These include products such as the Normalized Difference Vegetation Index (NDVI) produced from NOAA AVHRR data, land cover/land use classification data sets, precipitation, soil moisture, soil chemistry, surface temperature, and Digital Elevation Model (DEM) data sets. It is anticipated that in the future new instruments in the EOS program will make the critically important development of new products and improved algorithms possible.

This example demonstrates how research into a complex global change question requires the use of many kinds of data collected from many sources over a long time period.

4.4.3. Is the Earth Entering a Period of Extreme Climate Events?

To ascertain whether or not the Earth is entering a period of extreme climate events, it is necessary to determine whether or not hydrologic events are being amplified. Follow-up questions include: if hydrologic events are, in fact, being amplified, why? at what spatial and temporal scales can change be detected? are some regions more susceptible to change than others? what can be expected in the future? to what extent can such changes be attributed to human activities? Answering these questions will result in better understanding of the impact of climate variability on water and other natural resources, and on the nation's economy as well.

Developing answers to these questions requires long-term monitoring and rigorous investigation of many variables reflecting the interactions between water and energy fluxes at the land surface (i.e., precipitation, runoff, soil moisture, soil evaporation, and evapotranspiration). For example, changes in precipitation patterns (both in time and space), and the impact of these changes on the hydrologic cycle must be documented and understood. Similarly, long-term records must be examined to determine if droughts are becoming more common, if and where the severity of droughts is increasing, if the timing and magnitude of snow-accumulation and snow-melt on-set is changing, and if changes are taking place in regional vegetation cover. Global coverage of many state-variables over the longest possible time series is needed to understand the natural variability of the hydrological cycle so that deviations from the norm, such as a genuine progressive amplification of the cycle, can be detected with confidence.

While research efforts to derive some hydrologically relevant variables from remotely sensed data (e.g. soil-moisture) continue, other variables are currently available and need to be archived and documented. Examples of such existing products include fields of precipitation amounts, rates, and frequency (developed from GOES, SSM/I (Special Sensor Microwave Imager), NEXRAD and Surface Observations), temperature fields, atmospheric water vapor fields, snow cover extent, snowfall change, snow water equivalent, snow melt on-set, glacier mass balance / run-off, radiation balance products (including total incoming radiation, ground-leaving radiation, reflected and emitted). Also needed are various ancillary data sets, such as DEM, stream flow, water use/irrigation, ground water level, permafrost extent, and time-series of land cover classification. Because the ability to use these data sets in combination is essential to conduct effective investigations, archiving efforts must include facilitating the creation of such combinations of products. Finally, improvements are needed in the accuracy and resolution of the parameters contained in all of the existing data sets, as well as new soil moisture data sets.

In all cases, a careful and consistent documentation of algorithm changes and information about sensor characteristics and measurements is critical to correctly processing the data and must be part of a long-term archive. A closer look at one important parameter, atmospheric water vapor, illustrates and emphasizes this point. Water vapor is the most important of the greenhouse gases in the overall control of climate, and its condensate, clouds, modulate the radiative energy transfer; thus changes in water vapor concentration have major effects on climate (Elliot, 1995). The increase in water vapor concentration is linked to an enhanced man-made greenhouse effect.

This section raises what has become a major issue in the debate over the greenhouse effect, whether the amplitude or intensity of the hydrologic cycle increases as global temperatures increase. In Figure 17, the difference between the amount of water vapor in El Nino and La Nina years revealed by analysis of HIRS (High Resolution Infra-Red Sounder, flown aboard the NOAA series of polar orbiting satellites) data is quite apparent. However, tracking the long-term changes in water vapor has proved to be quite difficult. Progress is being made, for example as depicted in Figure 18, potential systematic biases have been identified in HIRS data due to differences in the filters used in successive HIRS instruments. These biases must be corrected for long-term climate studies.

This example again shows the need for combinations of many kinds of data to address science questions, but also emphasizes the importance of complete documentation, so that scientists are able to detect and compensate for problems like the systematic biases in HIRS data described above.

4.4.4 Are Ice Sheets Melting?

A possible consequence of global climate change in the form of an overall warming would be increases in sea level caused by the melting of portions of the polar ice sheets and mountain glaciers. Determination of whether or not such melting is in progress requires measurement of changes in the mass balance (between accumulation from snowfall and ablation from melting) of ice sheets and mountain glaciers. The kinds of questions that need to be addressed: Are changes taking place in the observable mass and extent of ice sheets and glaciers? Is sea level rising? Is atmospheric temperature changing over the ice sheets?

Examples of existing data products that are required for this study include: *in situ* ice sheet / glacier mass balance measurements, AVHRR and SAR (synthetic aperture radar) imagery, DEM, altimetry, iceberg calving records, ice motion, sea surface temperature, and surface air temperature. Examples of future products required include improvements in existing data products (e.g. improved high resolution visible satellite data and enhanced *in situ* monitoring of ice sheets and glaciers) and new products derived from laser altimetry, e.g. the GLAS (Geoscience Laser Altimeter System) to fly on IceSat in 2001.

These data must be used in combination to produce the highest quality information about ice sheet melting. For example, the top panel of Figure 19 shows data on the concentration of sea ice taken directly from satellites, which indicates nearly open water during the summer months around the north pole. But aircraft overflights and other data show that the sea ice has not melted, and it has been determined that the satellite instruments are sensing melt ponds that form on top of ice surfaces during the summer. A combined product of station, aircraft, and satellite data produced by the National Ice Center shown in the lower panel of Figure 19 provides a more representative picture.

This example shows that the ability to provide access to an integrated set of data compiled from a variety of sources is a critical function of a long-term archive.

4.4.5 Are Long-Term Changes Occurring in the Upper Stratosphere?

Currently, it is very difficult to accurately assess long-term changes occurring in the upper stratosphere, above the 10 millibar level. A major science need is for a better combined data set for the upper stratosphere. The NCAR/NCEP reanalysis data set (a combination of seven major *in situ* and satellite data sets covering in some cases a 50 year period of record) extends upward only to 10 millibars. A similar reanalysis data set extending up at least to the stratopause is needed.

There are several products which include satellite-based measurements of stratospheric quantities that should be continued and maintained. These include the Solar Backscatter Ultra-Violet instrument (SBUV) ozone profile measurements (which reach the upper stratosphere), MSU channel 4 weighted mean temperatures (lower stratosphere), and the TOMS ozone column measurements (lower stratosphere). Continued intercalibration analysis and combination of measurements of this type to produce a reliable long-term record for archival is needed.

This example points to an opportunity for an archive to work with scientists and data providers to produce a long-term, combined and intercalibrated data set needed to address a significant science problem.

4.4.6 Are There Observable Fluctuations / Trends in Solar Irradiance?

Solar spectral irradiance, and especially inter-cycle changes, is a potentially important climate forcing factor. Spectrally integrated solar irradiance has been monitored for about fifteen years, showing a decline of about 0.1% between 1979 and 1986, followed by at least a partial recovery (*Hansen et al.*, 1995). Solar variability of a few tenths of a percent could cause a global temperature change of the magnitude of the observed cooling between 1940 and 1970, and it has been suggested that the sun may be responsible for the warming trend of the past century (*Friis-Christensen and Lassen*, 1991). It is consequently important to monitor solar irradiance, and important to monitor the spectral distribution of any observed changes, because the forcing effect varies strongly with the altitude where the radiance is absorbed. Because of its importance as a climate forcing factor, spectral solar irradiance is a fundamental input to GCM (General Circulation Model) simulations of long-term climate change. For these reasons solar spectral irradiance data should have a high priority for data quality studies, intercalibration of data from separate instruments, and archival.

Currently, scientists still depend mainly on proxies for solar UV variability rather than direct measurements. Future measurements will probably be obtained by instruments similar to those on the UARS satellite. Careful intercalibration of measurements is needed to allow studies of inter-cycle solar spectral irradiance changes.

This example reinforces the need for complete documentation of data sets, mandatory for the kind of careful intercalibration required to allow detection of subtle changes over long time periods over which a number of instruments were used.

4.5 Summary of Lessons Learned

A fundamental lesson learned is that long-term archiving will be of enormous benefit to science, and thus to the nation through the results of research it enables, as long as great care is taken to preserve (in addition to derived products) low-level data (Level 1), with complete documentation, in ways that facilitate maintenance and user access. In particular, archiving of Level 1 data with complete documentation will make possible the successful generation of new geophysical products over the history of available data. In addition, storing data in a form that facilitates multiple reprocessing, inevitable as new product algorithms are developed and refined, saves time and expense, and facilitates use.

This lesson is supported by many of the examples cited above, including tropospheric temperature trends from long-term Level 1 MSU data (4.4.1), retrospectively finding and correcting for biases in long-term Level 1 HIRS data (4.4.3), ocean front density from SST's from long-term Level 1 AVHRR data (4.1.4), ozone analysis and trends from TOMS and other data (4.1.1), ocean topography from long-term Level 1 altimeter data (4.1.6), Pathfinder and other products from AVHRR (4.2.1), various products from OLS data (4.2.2), upper-tropospheric humidity from long-term Level 1 HIRS data (4.2.3), radiances as model inputs, from long-term Level 1 TOVS data (4.2.4), and geophysical products from long-term Level 1 SSM/I and SMMR data (4.2.5).

Another equally if not more fundamental lesson is the absolute importance of preserving complete documentation along with the data. Every example of successful reprocessing of archived data to produce improved or new geophysical products depends on documentation. In addition to the examples mentioned above, perhaps especially the ozone trends and AVHRR Pathfinder examples, this lesson is supported by trends in solar irradiance (section 4.4.6), global near-surface temperature trend analysis (4.1.3), snow cover trend analysis (4.1.5), and precipitation trend analysis (4.1.2).

Another fundamental lesson learned is that complex science problems associated with climate research will require scientists to use combinations of data sets (and their documentation) from many sources, satellite and *in situ*. Facilitating the use of data sets in combination, and intercalibration and intercomparison of data sets, will maximize support to researchers pursuing many such science problems. This lesson is supported in particular by these examples, study of changes in biospheric productivity (section 4.4.2), study of changes in the hydrologic cycle (4.4.3), study of ice sheet melting (4.4.4), global near-surface temperature trend analysis (4.1.3), and ozone trend analysis (4.1.1).

The example of the study of long-term changes in the upper stratosphere (4.4.5) highlights the key role that an archive can play if it seizes the opportunity to work in partnership with scientists and data providers to produce carefully calibrated, documented, integrated data sets needed for climate research. The recent NCAR - NCEP re-analysis effort is seen as a valuable precedent.

The following are additional specific lessons learned derived mainly from the AVHRR Level 1 (section 4.2.1), NEXRAD (4.3.1), GOES (4.3.2), and Seasat (4.3.3) examples:

1. Life cycle data management is needed, beginning at the planning and analysis stage, continuing through development and implementation, and should involve researchers and data managers as well as data collectors and primary users (operational or near-term research users).

2. Life cycle data management must be an ongoing process - it can not stop when an initial implementation is made, even if the initial implementation is the best that can be done at the time. Planning for and accomplishing migration to a system offering improved user access can be as important as migrating to a new archive medium.

3. Data managers have to be more aggressive, especially in the early stages of development. They must ensure that data management implications of system alternatives are fully explored from a research user perspective as well as that of the immediate operational user (in a case like NEXRAD) or the immediate principal investigator / research team (in the case of a research data collection effort, whether *in situ* or remote sensing).

4. Ready and responsive access to archived data is a key factor in securing its preservation - only through its use is the science value of data proven and the expense of preserving it justified. The LTA program must regard promoting use of the data as a part of its data preservation strategy.

5. Archive media must be chosen with an eye toward future migration as well as meeting current and projected user access requirements.

6. The LTA program must strive to maximize the use of standards based, ideally vendor independent technology, especially for data storage. The use of a non-standard or vendor dependent approach should be accepted only if a net advantage over a time period on the order of ten years is clear after a thorough evaluation of all aspects of the trade-offs involved. In either case, costs for migration to a different storage technology should be put into the program's budget from the outset.

A last fundamental lesson learned is that the lack of a program focus for long-term archiving, with active engagement of the science community, has left decisions on whether and/or how to archive data sets to be handled case by case by agencies within operational or research programs whose primary goals are near-term, short-term, or even only real-time. A prime example of this is the AVHRR and TOVS case discussed in section 4.2.1. Long-term archiving and concern for future use and value of data sets beyond immediate needs has often been nothing more than a low priority after thought, and always at a disadvantage in the competition for resources. The availability of a fully capable approach to long-term archiving has not guaranteed success in obtaining the needed resources - especially if a less expensive solution is seen as meeting short term needs (e.g. NEXRAD, section 4.3.1).

5.0 Essential Functions and Characteristics of the LTA Program

This section discusses the essential functions and characteristics of the LTA program. As noted in section 2.0 above, the term "program" implies an organized approach, documented in a program plan, supported by a dedicated line item budget, with people held responsible for carrying out the program and accountable for its success or failure. In practice the functions discussed in this section would be performed by the operating elements of the program (federal data centers, cooperating science groups, etc.).

The workshop developed a statement of general functions, which were then broken down into a set of graded specific functions. This section presents the general functions, discusses the grading, and then describes the specific functions. In considering these, the workshop bore in mind lessons learned from experiences of existing archives summarized in section 4.5 above, seeking to reinforce what has proven successful, and to build in safeguards against what has not been successful.

5.1 High Level LTA Program Functions

At the highest level, the functions of the LTA program are to:

1. Ensure that mechanisms are in place to engage the science community in decisions concerning which data/products are to be included in or excluded from the long-term archive and to review the contents of the long-term archive for disposition;

2. Ingest and verify the content and quality of incoming data, derived products, *and their documentation;*

3. Preserve and maintain the basic storage of essential long-term Level 1 data and derived products and full documentation – including media monitoring and refreshment and migration of data to new media and storage/access systems as needed to meet user needs;

4. Perform and sponsor some reprocessing and generation of new derived products;

5. Provide open access to the data - distribute the data to the maximum number of users at the minimum cost to the user, including informing potential users of available data, products, and services.

Here and in the sections below, the descriptions of functions and characteristics should be thought of as science requirements for the LTA program – i.e. specific science requirements are expressed as a set of functions or characteristics the program and its operating elements should have.

One overarching fact that must be allowed for in planning the budget for the LTA program is that the archive holdings will grow significantly with increasing volumes of current and new data

sets. Some cost factors will scale with sheer data volume, others will scale with the number of data sets or derived products; and although there may be some technological off-sets, the increasing volume and number of data sets must be explicitly allowed for in the preparation of budget projections.

Scientists at the workshop also emphasized the importance of considering both the costs of getting a data set safely written to a given storage medium, and the cost of reading the data back, especially of accessing a long time series or the full record for a data set. Simply adopting an inexpensive medium on which to write data can be a terribly false economy and highly counter-productive if the access costs and time required effectively prevent the kind of research (examination of long time series) essential for the study of climate variability and trends.

The description of science requirements and derived data management requirements that follows below is consistent with material included in Appendix C about NASA's Earth Science Program, the IPCC recommendations, recommendations from the Workshop on Long-Term Climate Monitoring, the Data Management for Global Change Research Policy Statements and key points made in the 1995 NRC study "Preserving Data on Our Physical Universe".

5.2 Detailed LTA Program Functions and Characteristics

On the basis of the high-level functions discussed above, this section defines seven detailed functions: user involvement, data administration, documentation, data ingest and verification, data preservation and maintenance, data processing and reprocessing, and finally data access and user support.

The scientists at the workshop developed a rough scale for scoring the importance of specific functions. They were classed as either "**essential**", in which case they were regarded as required no matter what they would cost to implement, or "**desirable**", in which case they should be provided if the cost to do so would be low, or if an analysis shows that the benefits would be very high in relation to cost. The goal of the rating, in addition to a statement of relative importance, is to support the development of cost assumptions as part of the process of producing a budget request.

While marking some functions as "**essential**", scientists at the workshop recognized that budgets will be constrained, that first class treatment will be affordable only for vital data sets, that the science community must set priorities, and that concern for data should be every scientist's second priority after his or her own research. This means that there will be gradations of performance or 'level of service' within some of these functions, and that while an "**essential**" function must be provided for all data sets, the actual implementation and consequently some aspects of performance may vary in accordance with the relative priority of the data set. For example, an infrequently used data set might be stored off-line on a medium/equipment combination with a lower transfer rate and a longer access time, than a data set maintained in near-line storage and more rapidly accessible – but the requirement to preserve and maintain both data sets *and their documentation* would be held with the same full force.

5.2.1 User Involvement

It is **essential** that scientists be involved with all aspects of the LTA program. The purposes for the LTA program can only be successfully achieved, and the vision fulfilled, if the program meets the needs of the science community. This in turn can only be achieved through ongoing involvement and interaction between the community and the LTA program. The burden rests with the program to implement mechanisms and approaches to securing this involvement and to ensure responsiveness to science priorities and advice. The program must be a working partnership between data managers and scientists within the archive program and other scientists who use the program's services. The program must include experienced, knowledgeable scientists who are involved in exercising the data and generating products, performing quality assessment / quality control of incoming data, products, and documentation, of internally generated products and documentation, and products and services provided to users. Scientists within and outside of the program must be engaged in planning strategic direction for the program.

The requirement for user involvement is an overarching requirement that touches all areas of the LTA program. In particular, the discussions of data administration (section 5.2.2 below) and data access and user support (section 5.2.7 below) will elaborate on user involvement with those LTA functions.

5.2.2 Data Administration

This function is concerned with administering the content of the long-term archive. It includes planning / deciding what data should be included in the long-term archive, excluded from it, or removed from it, and what the criteria will be for making such decisions, and how the decision making process will work.

It is **essential** that the LTA program preserve key long-term data and information, including the definitive version of the EOS Level 1 data and derived products, NEXRAD, GOES, NPOESS and any other data sets / products needed to interpret and use them, and any other satellite and *in situ* data sets needed for calibration, and validation, data to be used in reprocessing and production of new higher-level products, and essential improved/reprocessed products.

It follows that it is equally **essential** that the program establish a process for deciding which products to include within or exclude from, or remove from, the long-term archive, and that this process be driven by science priorities and scientific assessment of the products in question, and therefore that scientists be actively engaged in this process, in defining procedures, setting criteria and making decisions. Some products may come to the LTA program already approved (e.g. EOS standard products peer reviewed by the EOS Investigator Working Group or products for which NOAA has an operational requirement for archiving).

In establishing this process for data administration, the LTA program must follow the guiding principle of keeping things as simple as possible; it must ensure that this process does not become cumbersome, rigid, or overly bureaucratic, and that it works simply, and doesn't get bogged down. Information will be fed into this process from a number of sources. Routine

operating functions of the LTA program will be a source of information about product quality (including feedback from users) and usage.

It is **desirable** that the data managers within the LTA program, in collaboration with scientists, identify, locate and recommend for inclusion new data sets and/or derived products, including ancillary data sets required for intercomparison and validation, and also recommend data sets and products for removal.

5.2.3 Documentation

Documentation (which includes everything meant by the term "metadata") of an archived data set or derived product is the complete package of information describing the data or product sufficient to ensure its usefulness to the scientist, and subsequent generations of scientists, with no prior knowledge of the data set or product. Documentation includes (as appropriate for the level of the data set or product in question) but is not limited to the following:

1. Instrument / sensor characteristics including pre-flight or pre-operational performance measurements (e.g. spectral response, noise characteristics, etc.);
2. Instrument / sensor calibration data and method;
3. Processing algorithms and their scientific basis, including complete description of any sampling or mapping algorithm used in the creation of the product (e.g. contained in peer reviewed papers, in some cases supplemented by thematic information introducing the data set or product to scientists unfamiliar with it);
4. Complete information on any ancillary data or other data sets used in generation or calibration of the data set or derived product;
5. Processing history including versions of processing source code corresponding to versions of the data set or derived product held in the archive;
6. Quality assessment information;
7. Validation record, including identification of validation data sets;
8. Data structure and format, with definition of all parameters and fields;
9. In the case of earth-based data, station location and any changes in location, instrumentation, controlling agency, surrounding land use and other factors which could influence the long-term record;
10. A bibliography of pertinent Technical Notes and articles, including refereed publications reporting on research using the data set;
11. Information received back from users of the data set or product.

A data set or derived product can be thought of as a type comprised of a series or collection of real instances of the type. For example, a global field of 12 hourly, synoptic 500mb temperatures and dew points on a 5 degree latitude/longitude grid could be a product, and a series or collection of 7,312 of these fields might exist for 0Z and 12Z for every day for a ten year period starting with January 1, 1980. Some elements of documentation apply to the data set or product as a type or collection, i.e. are common to all instances (such as parameter definitions or gridding method), some will vary instance by instance, (such as dynamic measurement of calibration parameters, time tags, geo-location, and quality flags).

It was deemed **most essential** that the LTA program ensure that data sets and products in the archived are accompanied by complete, comprehensive, and accurate documentation. This was seen as critical for future access and usability and for meeting the "20 year test" guiding principle. Poor documentation was seen as destroying the value of a data set or product.

It is **essential** that the LTA program ensure that documentation (which must include source code of processing software) of a data set or product is complete and suitable for use by scientists who are not experts with that data set or product and indeed who may be entirely without prior knowledge of it. To that end, the scientists at the workshop saw a peer review process as **essential,** which would include scientists unfamiliar with the sensor, data set, or product reviewing the documentation on behalf of future generations of scientists. It was recognized that a significant pitfall for documentation is what the instrument scientist or scientist who develops a product regards as "common knowledge" and thus either fails to include in the documentation or alludes to without explanation. That "common knowledge" may not be shared by anyone else at the time, let alone future generations of scientists. Review by scientists unfamiliar with the data set or product would help to catch such deficiencies in the documentation.

While documentation was recognized as a shared responsibility of instrument/sensor developers, science data creators, system designers, and data managers, the LTA program must exercise its responsibility proactively to establish appropriate relationships with science data producers (both operational and research) to begin a process of active, ongoing cooperation in the development of documentation. The onus is on the LTA program to ensure that its holdings are completely documented.

5.2.4 Data Ingest and Verification

Ingest includes taking into the archive new data sets and products and/or new versions of existing data sets or products or replacements of previous versions, as approved and directed by the data administration function discussed above.

It is **essential** that the LTA program verify the integrity and quality of data and derived products and associated documentation as it is ingested into the archive.

A data set or product being ingested into the archive, and its overall data set level or product level documentation, would have been approved as discussed above. In addition, each instance of the series of instances comprising the data set as a whole, or each instance of a product, must be verified and the instance level documentation must be verified as the instances are ingested and added to the archive. This is necessary whether the instances come to the archive in one bulk collection or as an operational or continuing stream.

It is **essential** that the operational LTA program be capable of ingesting data and/or derived products and their documentation from research environments, because at the present time, and perhaps more frequently in the future, new data collection systems will be operated in research environments and new or improved products will also be generated in those environments. This places a special burden on the LTA program to proactively reach out to the research data source and develop the needed agreements and procedures, assist in planning documentation, etc., as

well as agree on the means and schedule by which the actual transfer of data will occur. Resources to support both the long-term archive and the data producer to complete the planning for and execution of the transition are required.

5.2.5 Data Preservation and Maintenance

The data administration section above discussed maintenance of the archive's base of data and information from the point of view of science content – which data sets and products should be included or excluded, i.e. which should reside within the archive. This section considers preservation and maintenance of the data and information while it is resident within the archive. Preservation and maintenance of data holdings, including ensuring integrity and quality of the data and associated documentation, is an **essential** function of the long-term archive. Extension of maintenance to include updating of data set or product documentation with user comments was considered **desirable.**

It is **essential** that the LTA program perform integrity checks on archive media between data migrations – e.g. by random sampling, tracking hardware error reports, with a strategy for hardware/media error control.

It is **essential** that the LTA program develop and maintain a multi-year data migration plan, in which at least the following factors driving migration are considered: deteriorating media (based on information developed from user accesses and data integrity checks) technically obsolete media, and obsolescing access systems. This must include planning and budgeting for not just media refreshment, but the replacement of storage and access equipment as well. This adds both expense and savings – expense in buying new equipment, and in many cases reformatting of the data (which requires that the data files be well understood, lest the data are corrupted in the process) and savings in floor space required, media handling, and equipment cost, since new equipment and media invariably hold much more data than old ones of the same volume, and are less expensive to maintain.

The National Archives and Records Administration (NARA) provides guidance to federal agencies concerning storage, preservation, and disposition of "federal records", a generic term embracing, for NOAA and NASA, data, derived products, and documentation. Compliance with NARA regulations was deemed to be **desirable** by scientists at the workshop. Literal compliance would protect the LTA program from some degree of political vulnerability in the event of a serious data loss, but the LTA program would be held accountable by the science community for a serious data loss regardless of compliance with NARA regulations (hence the "**essential**" preservation requirement noted above).

Literal compliance with NARA regulations could also lock the program into a very conservative technological approach that would add cost and compromise performance, again undesirable from the point of view of the science community. It was noted that there would likely be no archive of Level 1 AVHRR and TOVS data, GOES data, low-level NEXRAD data, and doubtless other important data sets if literal compliance were to have been an "all or nothing" choice given the resource constraints faced by NOAA and NASA over the years.

The scientists at the workshop favored costing out both fully compliant and reasonable compromise approaches, working with NARA to secure approval of storage technology that would meet performance requirements within cost constraints, and otherwise to use compromise approaches where cost / performance problems proved to be severe. It was noted, for example, that careful monitoring and planned frequent media refresh well within documented media lifetime would meet the spirit of the NARA regulations, or that using NARA compliant media for backup of 'working copies' on high performance media was an option to consider. Operation of redundant systems, i.e. offering a full suite of data and information services from two separate locations was offered as a possible approach.

In sum, preservation of data and information is a top priority requirement of both the science community and NARA, but the LTA program needs the flexibility to meet this requirement in ways that allow it to provide the best possible suite of services to the science community within a constrained budget environment.

5.2.6 Data Processing / Reprocessing

It is **essential** that exercise of data to produce new products and/or new versions of old products be performed by the LTA program.

The processing / reprocessing exercises should be done both by scientists in-house to the LTA program and by scientists in the broader community. The LTA program should fund competitively selected proposals, with a budget in the range of 1% to 10% of the program's operating budget. The only way to bring back to the program the expertise gained in processes done by scientists in the broader community is for at least one person from the program to be actively involved with any such reprocessing.

The purposes of processing / reprocessing exercises are multiple:

1. Validate data and product documentation.

2. Identify and resolve problems in the data; inevitably data users find problems in the data, its storage, or documentation, which the archive can either fix or incorporate into the information describing the data.

3. Provide opportunities to scientists within the LTA program to pursue science interests, recognizing the need to maintain their involvement for the health of the program. A less obvious but important effect of reprocessing is to interest data center personnel. When data center personnel are prohibited from involvement in the generation of higher level products the better ones tend to leave, and all become disinterested in the data they manage.

4. Produce new products or new versions of old products that are of value to the science community – reprocessing of low-level data is likely to be done a few times by a very small number of users with the necessary interest, expertise, and capability, but it can yield new or improved high-level products of interest to many users.

5. Provide an opportunity to rethink and reorganize how the data are stored to take into account user access needs as well as accommodate new storage and access technology.

6. Increase data longevity - exercising the data by interested research users, both inside or outside the data center, is by far the best way to increase data longevity.

Processing / reprocessing efforts involving some data sets will take place in the normal course of affairs. The objective of the LTA program's exercises should be to ensure that data sets that are not otherwise exercised, get exercised.

Data migrations should be taken as opportunities for processing / reprocessing – a competitive request for proposals process can be used to determine if there are worthwhile products to produce in the course of a data migration.

5.2.7 Data Access and User Support

It is **essential** that the LTA program provide the next and subsequent generations of scientists with appropriate access to, and facilitate their use of, its holdings. "Access" includes a data set / product search and order function, the ability to deliver data and/or products and supporting information (documentation) on suitable media or electronically, and choices of data structure and format, user options such as subsetting, etc., that facilitate access and use (for example by minimizing the volume of data that the user has to contend with). A vital part of "access" is user support staff knowledgeable about the data, willing and able to help users identify, obtain, and use the data sets and products they need, including making referrals to other sources of data. "Appropriate" means first in a manner responsive to user needs as discussed in the next paragraph, but also with full consideration given to the guiding principle of simplicity. Most users' only experience with the LTA program will be with its access and user support functions – at all stages of design, implementation, and operation the question should always be asked, "is this the best we can do to make this service easy to use, simple for the user?".

It is **essential** that the LTA program provide data and information services that are responsive to needs of its users. While its services are available to anyone, the program should emphasize as **essential** meeting the needs (e.g. by taking steps as described below) of research scientists and users who perform assessments - users who pass on knowledge of what the data mean to people such as policy makers or business decision makers. Other users, beyond the science users, will have access to LTA program services, but taking additional steps to meet their particular needs is **desirable**, and is unlikely to be affordable. But many needs of other users, such as for documentation and straightforward means of access to data and products, are shared with the science community and will be met, albeit almost certainly without special tailoring for non-scientists.

Being responsive to user needs requires that the LTA program must know its user community, understand community needs, set out to meet those needs, and must have and use mechanisms to assess how well it actually does so, and then must seek to remedy shortcomings. Knowing the community and understanding its needs requires that the program be actively engaged with the

community, through use of mechanisms such as an active user support staff and a user advisory committee, attendance at user conferences, soliciting and collecting individual comments on services, data, products, etc. Setting out to meet user needs means designing services explicitly to that end, whether it be in data formats, options such as spatial/temporal subsetting, facilitating network access or in choice of distribution media, catalog content and search methods, improving response times, number and science/data background of user support staff, etc. Having mechanisms for assessing how well user needs are actually met involves collection of appropriate metrics and user feedback, and active discussion of these with a user advisory committee, and development of a clear picture of performance against needs that such a committee endorses as realistic. Finally, seeking to remedy shortcomings means that the program must be capable of accepting conclusions drawn from the analysis of performance against need and the judgement of the users, and working as needed to rectify problems.

The access functions of the LTA program should be designed to:

1. Enable the development of new or improved algorithms, especially those that rely on large and/or long-term data sets (e.g. organize/structure data in the LTA in a way that facilitates use for long time series test runs of new algorithms);

2. Support/enable reprocessing and generation of new products – for example by making the required data easily accessible by the same method used by the original principal investigators;

3. Enable ready and easy access to data sets needed for intercomparison and validation – offer options such as common gridding or formatting that would facilitate intercomparison by users, including comparisons between satellite and in-situ data sets (noting that satellite data are used to help work through the consequences to the long-term record of changes in in-situ instrumentation).

Design of access functions and features should not necessarily treat all data sets in exactly the same way. For example, ease of access (perhaps especially by network, and rapidly in times measured in minutes and hours to a day) should be emphasized for higher level products that will be needed by a larger number of applications oriented vs. data expert users, including users doing assessments as well as researchers. On the other hand, a much smaller number of users more expert in the details of the data will require access to larger volumes of lower level data – for them straightforward bulk delivery in a simple format on economical media in a reasonable time measured in days to a week is be appropriate.

Scientists at the workshop noted an emerging extension to the current access paradigm that the LTA program ought to consider. For example, data sets and products in the NCAR archive are held on-line to the NCAR computing facility, so that scientists at NCAR or remote users of the facility can develop and run applications software that directly access the NCAR archive without the scientist having to go through a separate process to obtain the needed data sets. Scientists speculated that the 'archive of the future' might include a generalization of this concept, an on-line data server accessed by user applications running on a user's workstation or an arbitrary

remote computing facility, with the access accomplished by the user's application in a manner transparent to the user.

Scientists at the workshop engaged NASA and NOAA managers in a discussion of data pricing policies. While there was general recognition that charges for data at some level were likely to be unavoidable, the following general principles were stressed:

1. If the data are not affordable by research users, they will not be used in research; so that if the vision and fundamental purposes of the LTA program are to be achieved, the cost of data and products to the research user must not become an insurmountable obstacle to research.

2. Prices charged for data and products should at most only cover incremental costs involved in actually filling the user requests ("COFUR - Cost of Filling User Request") and should not attempt to cover other portions of the LTA program budget.

3. Corollary to the second principle, future generations of users cannot be expected to pay for current costs – i.e. the cost of building and maintaining the archive can not be passed forward to future generations of users in the form of charges above the COFUR level.

The LTA program should establish a plan and process for setting and maintaining prices for data and products. The plan and process should be consistent with the above principles and should be a topic for open and full discussion with the science community. Any implementation should adhere to the simplicity principle.

6.0 Conclusion

Climate change and the consequent impacts and hazards on society is a very important aspect of the future. The best possible scientific understanding of these changes is essential for the economic well being and quality of life of the American public and the economic strength of the United States. The knowledge gathered from climate change research will help guide future public policy, business and personal decisions. The importance of understanding human impacts on our surroundings is increasing as American society becomes more urban and more dependent on robust infrastructure and production, transportation, and delivery of a wide range of products and services. Assessment of these impacts, possible only through greater scientific understanding of the Earth system, is crucial.

The scientific answers necessary for improved understanding of the Earth system can only be obtained through a wide variety of global-scale, long-term measurements. Instruments flown on satellites, carried aloft on balloons, flown on aircraft, deployed by ships, attached to buoys, and installed at fixed land based stations all contribute to global change research. Often operational collection systems, such as real-time weather prediction, can be used for purposes originally unintended and applied to global change research. Although a unique aspect of the EOS project is that it is designed specifically for global change research, in order to extend the record and to

cross-calibrate and validate the new measurements, researchers will have to make use of previously collected information. The EOS project will only experience maximum success if the data, derived products, and complete documentation of other projects are preserved and maintained in a carefully designed LTA program. With easy accessibility and well-planned support services, the LTA program will be of great benefit to the scientific community and the public.

A successful LTA program will only be achieved if:

- Existing long-term archiving activities for atmospheric and oceanographic data and derived products are coordinated and elevated to the status of a distinct program with a dedicated budget and specific stipulations for responsibility and coordination;

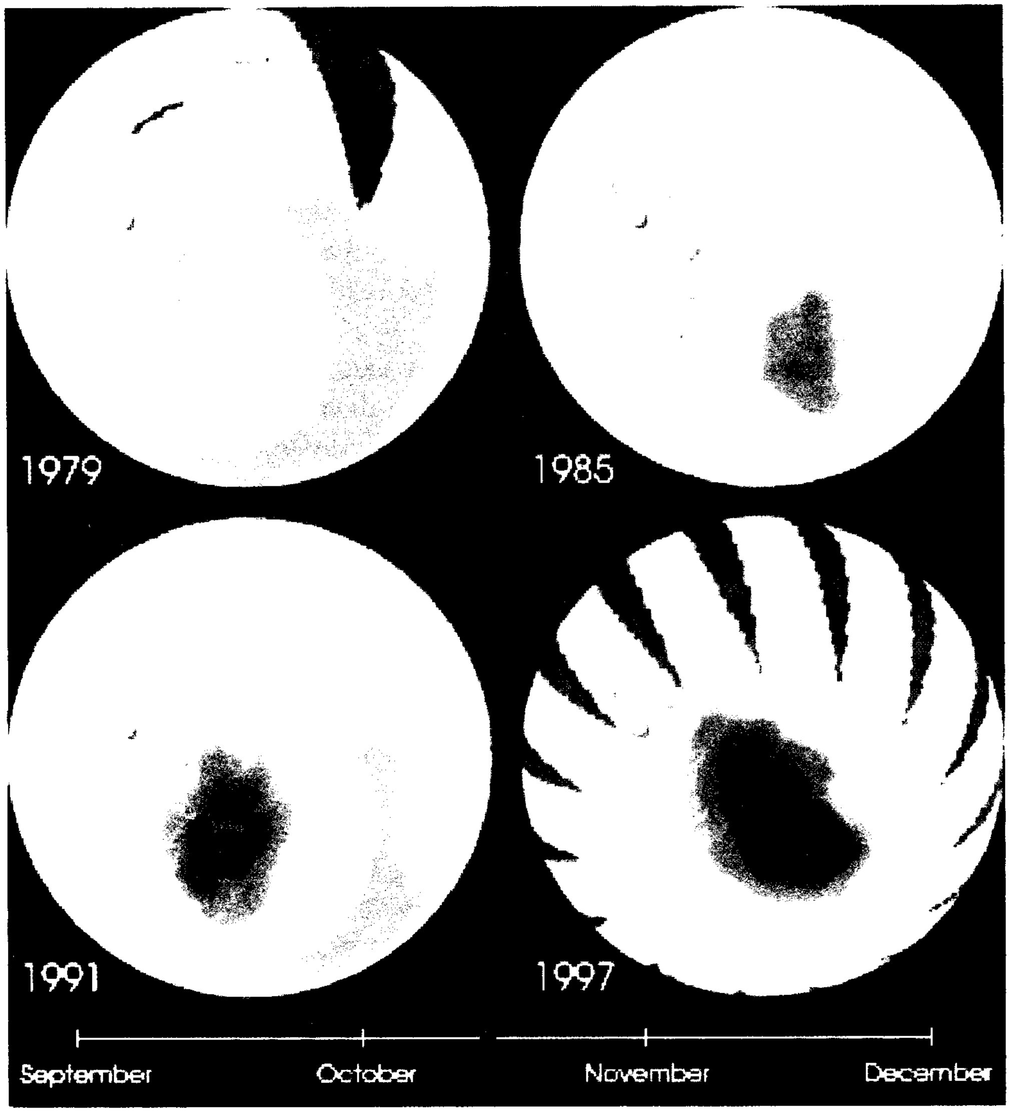
- The present scope of long-term archiving is extended to include NASA EOS and related NASA atmospheric and oceanographic data and derived products, NPOESS data and derived products, and fully effective handling of NOAA GOES and NEXRAD data sets;

- It embraces the vision and guiding principles arising from this workshop, commits to the purposes set forth there and reported above, and provides the recommended mix of "essential" and "desirable" functions in a cost-effective manner; and

- It fosters dedicated conservation of the nation's Earth science data and information resources and builds an effective and vigorous partnership between data managers and scientists.

The vision, guiding principles, purposes, functions and characteristics of the LTA program developed at the workshop and described in this document are fully consistent with the objectives and needs of the U.S. Global Change Research Program, NASA's Earth science research goals, the data management dimension of the International Panel on Climate Change's requirements for climate monitoring, and the recommendations of the Workshop on Long-Term Climate Monitoring.

Figure 1

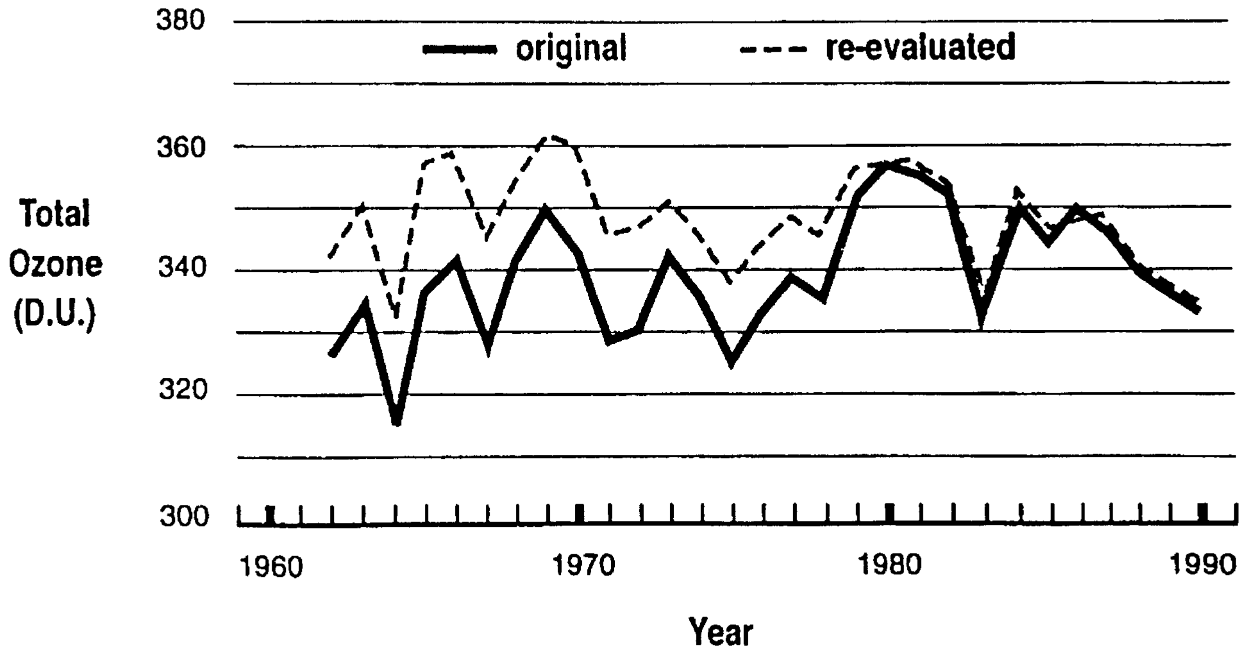


Antarctic Ozone Hole Progression

The four panels are for mid-October days, in 1979, 1985, 1991, and 1997. The first three images are from Nimbus-7 TOMS data and the fourth from Earth Probe TOMS data. The color scale runs from about 400 Dobson units (red) down to 100 Dobson units (deep blue) of total column ozone, so the growing blue area over Antarctica indicates the growing "ozone hole".

Source: Steve Kempler, Goddard Space Flight Center, DAAC From section 4.1.1

Figure 2



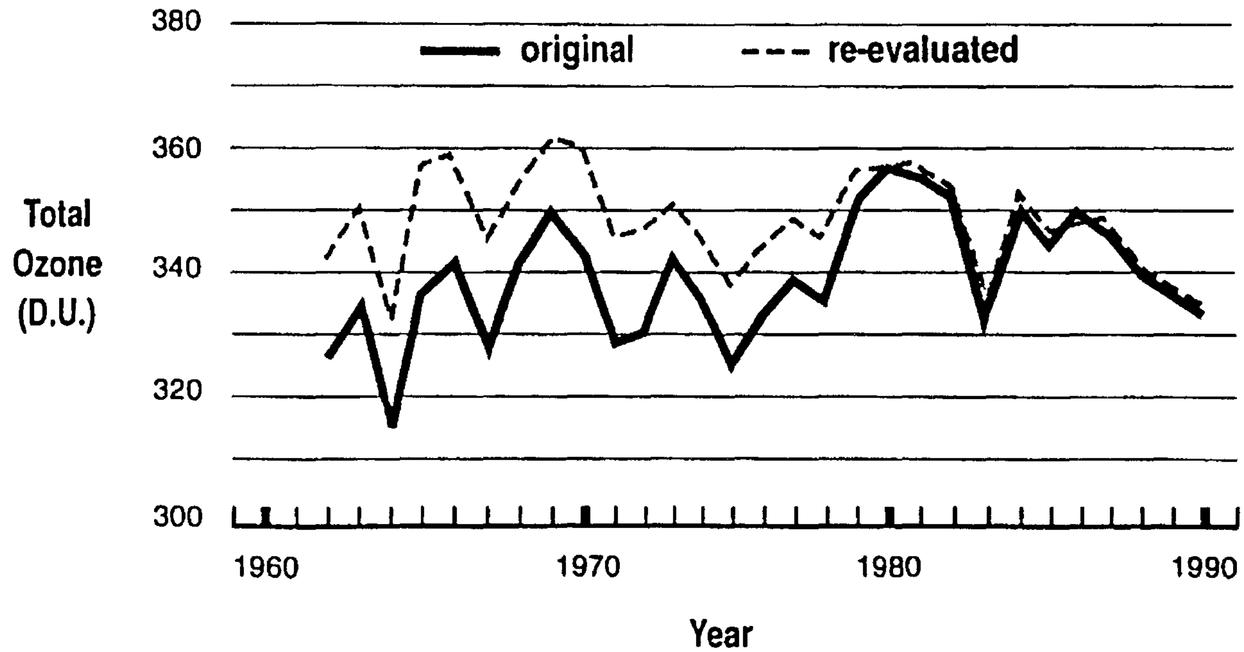
Annual Means of Total Ozone

Changes in total ozone, Hradrec Kralove, prior and subsequent to re-evaluation of calibration.

Source: Figure courtesy of W. Planet, NOAA in "Long-Term Climate Monitoring by the Global Climate Observing System" edited by Thomas R. Karl, p.61

From section 4.1.1

Figure 3

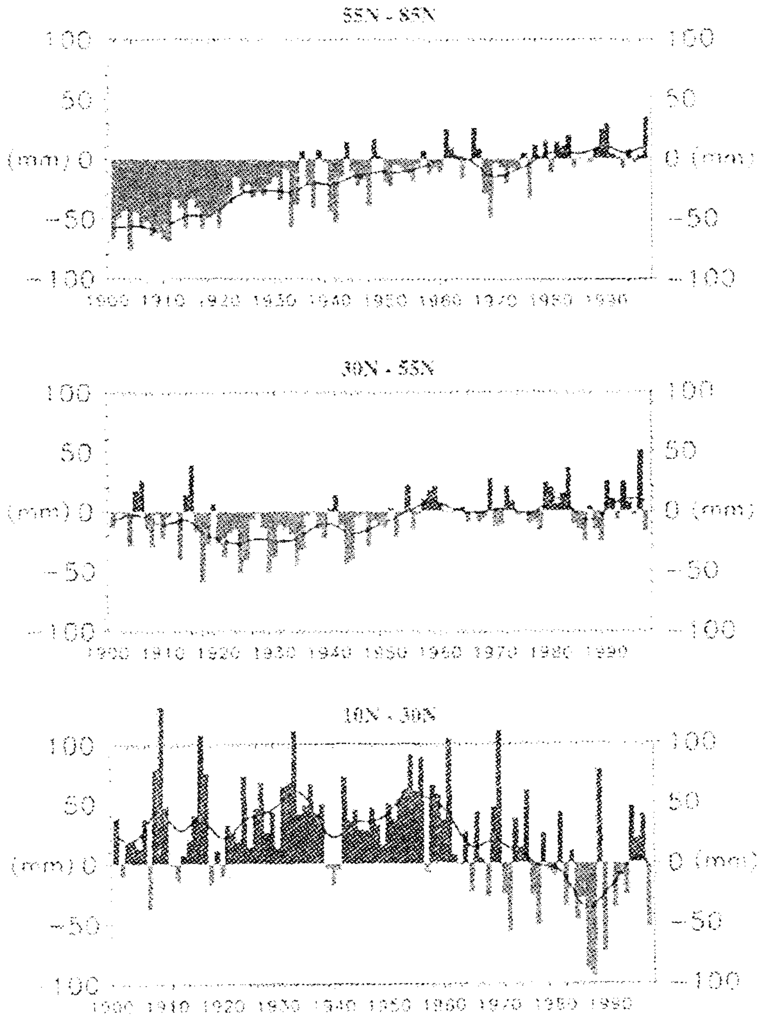


Instrument Comparison

Systematic differences of total ozone as measured by satellite SBUV2 and *in situ* Dobson measurements. SBUV2 data was processed by algorithms version 5.5 (NOAA-9) and version 6.0 (NOAA-11).

Source: Figure courtesy of K. Rao, NOAA in "Long-Term Climate Monitoring by the Global Climate Observing System" edited by Thomas R. Karl, p.61 From section 4.1.1

Figure 4

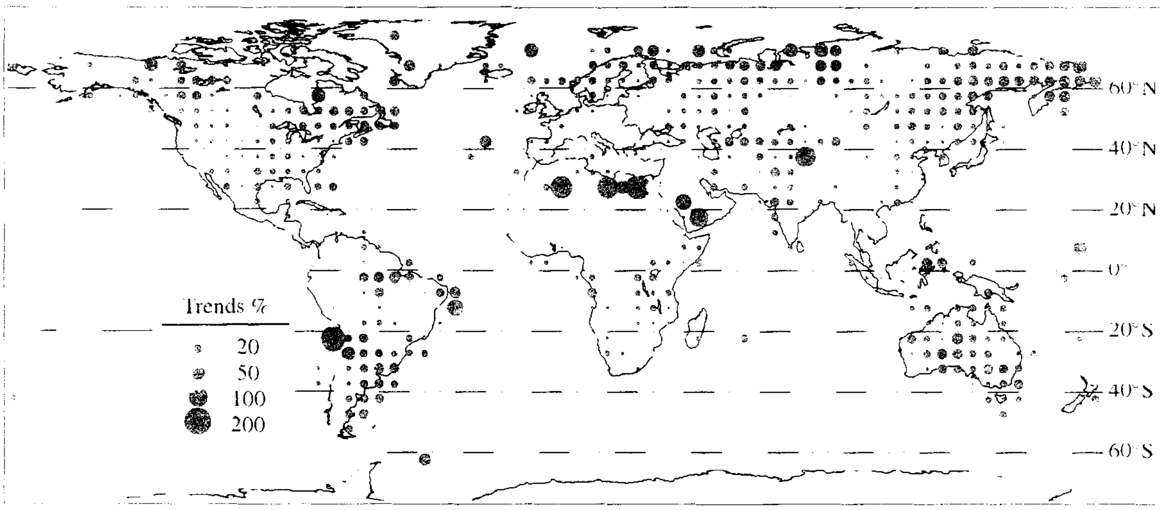


Observed Precipitation Changes

Annual anomalies of precipitation averaged across the global land areas. There has been a 1% increase in precipitation globally since the late 19^{th} century. Precipitation has increased in mid- to high-latitudes and decreased in the sub-tropics.

Source: Tom Karl, National Climatic Data Center, NOAA-NESDIS From section 4.1.2

Figure 5

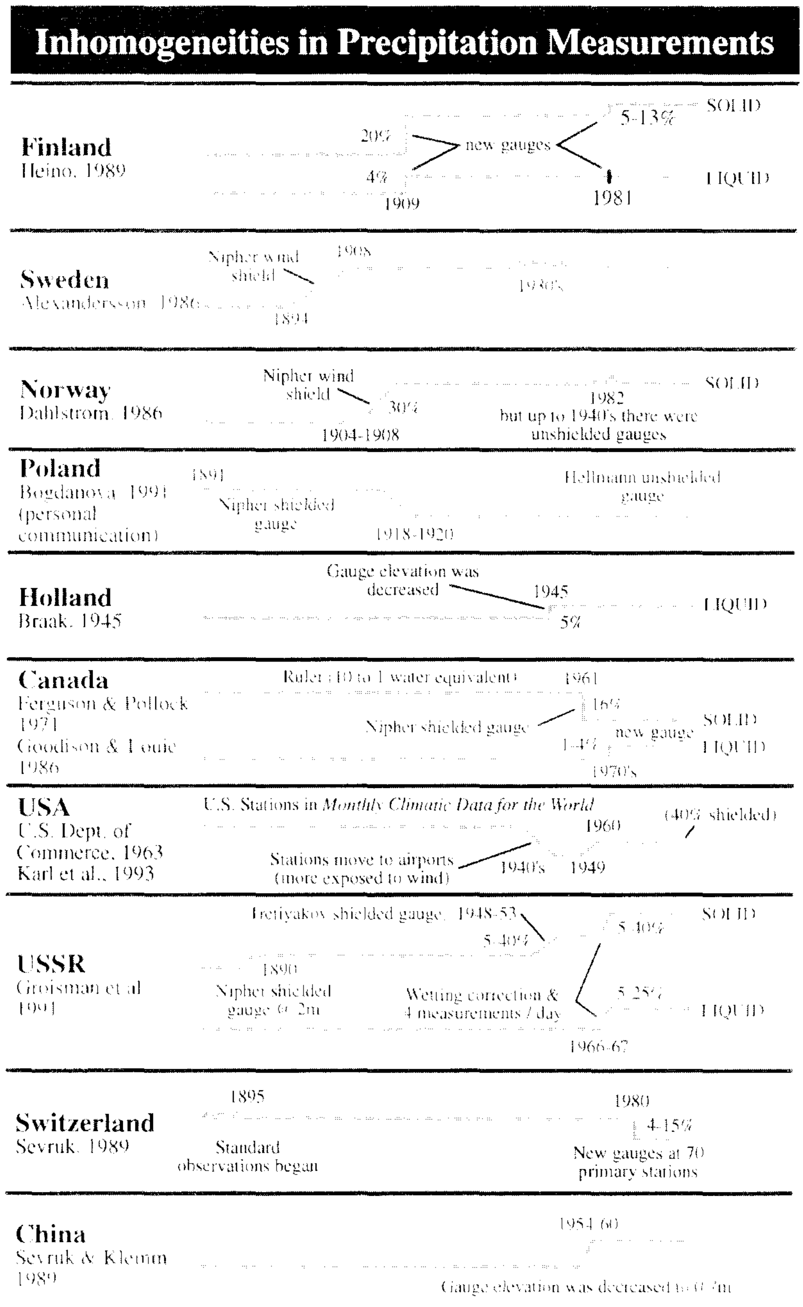


Mean Annual Trends of Precipitation (1901-1995)

The magnitude of the trends for each 5°x5° grid cell is given by the area of circle centered in each cell; brown circles reflect decreasing trends and green circles increasing trends. Trends are given in %/century and are expressed in percent relative to the 1961-1990 average precipitation.

Source: Tom Karl, National Climatic Data Center, NOAA-NESDIS From section 4.1.2

Figure 6

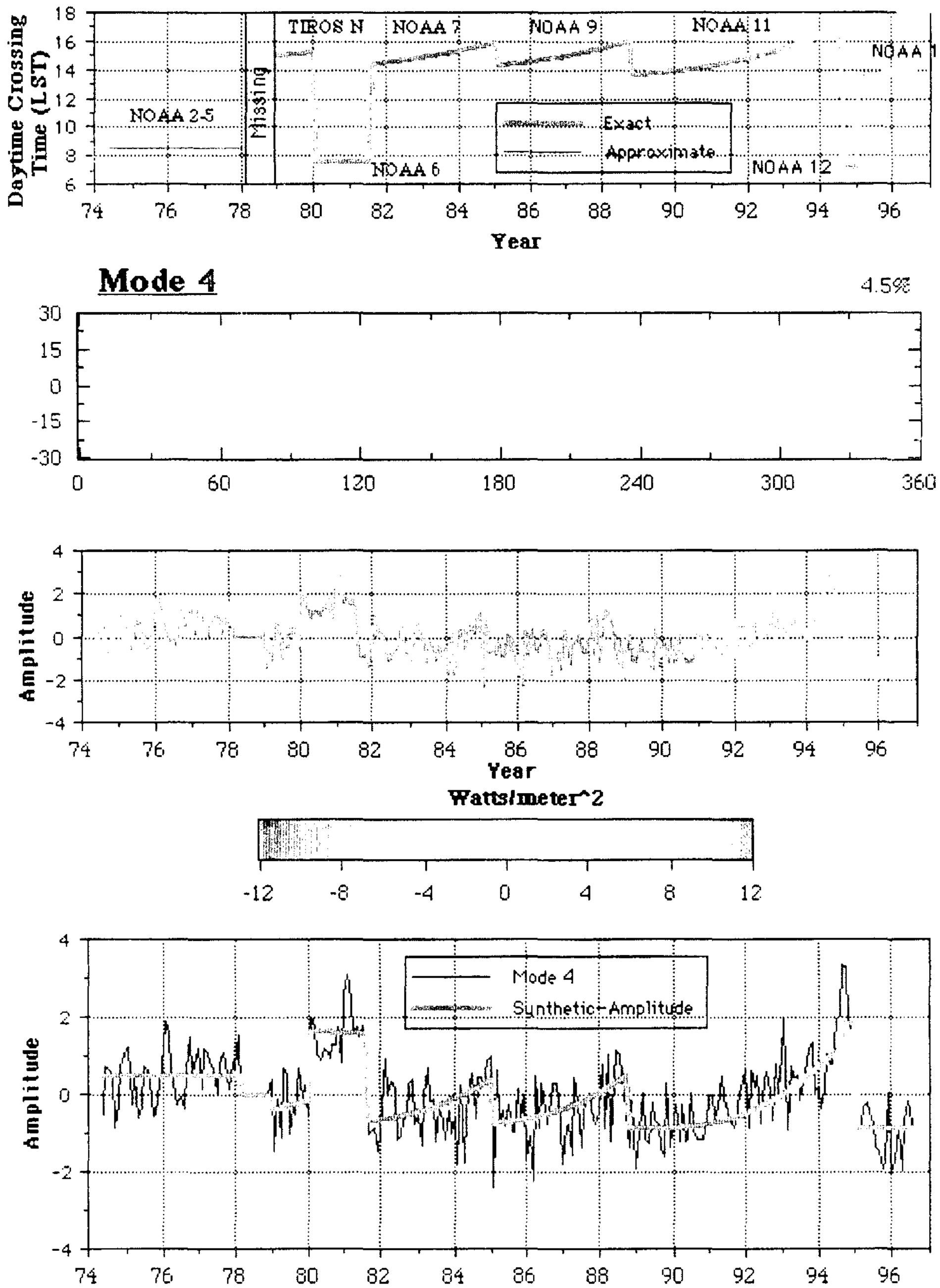


Inhomogeneities in Precipitation Measurements

Information on how the precipitation in situ observations were made, the changes made in observation techniques and instrumentation over time, and the variations in the precipitation measurements that resulted.

Source: Tom Karl, National Climatic Data Center, NOAA-NESDIS From section 4.1.2

Figure 7

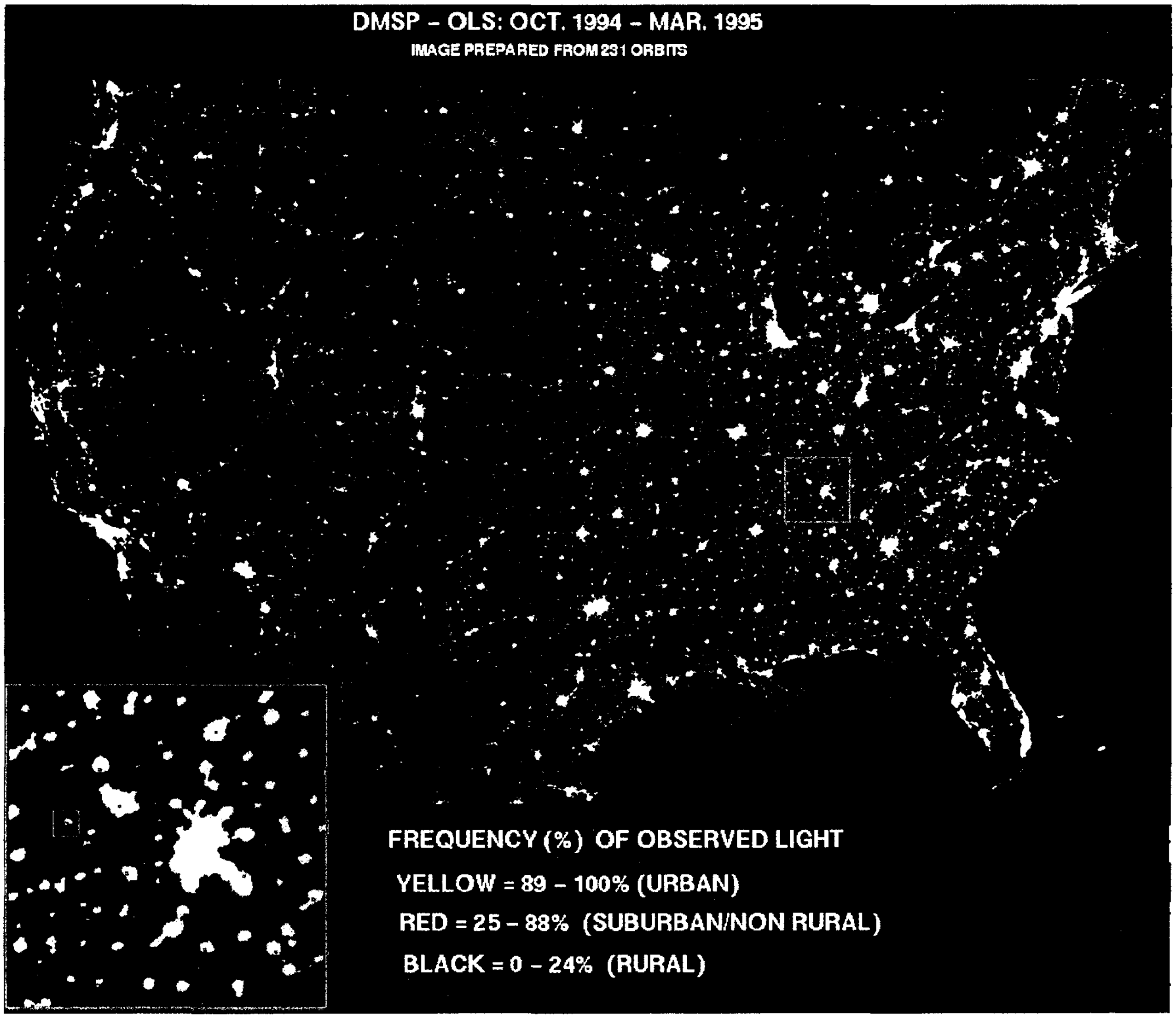


Orbital Shift and HIRS Data

The top panel shows the time-drift in the daytime equator crossing for NOAA's series of polar orbiting satellites. Changes in the time of day of the satellite overpass can produce false trends that look like climate change, but are in fact just artifacts of the satellites' orbital characteristics. Changes in the radiation values track the orbital drift of the satellites.

Source: Tom Karl, National Climatic Data Center, NOAA-NESDIS From section 4.1.2

Figure 8

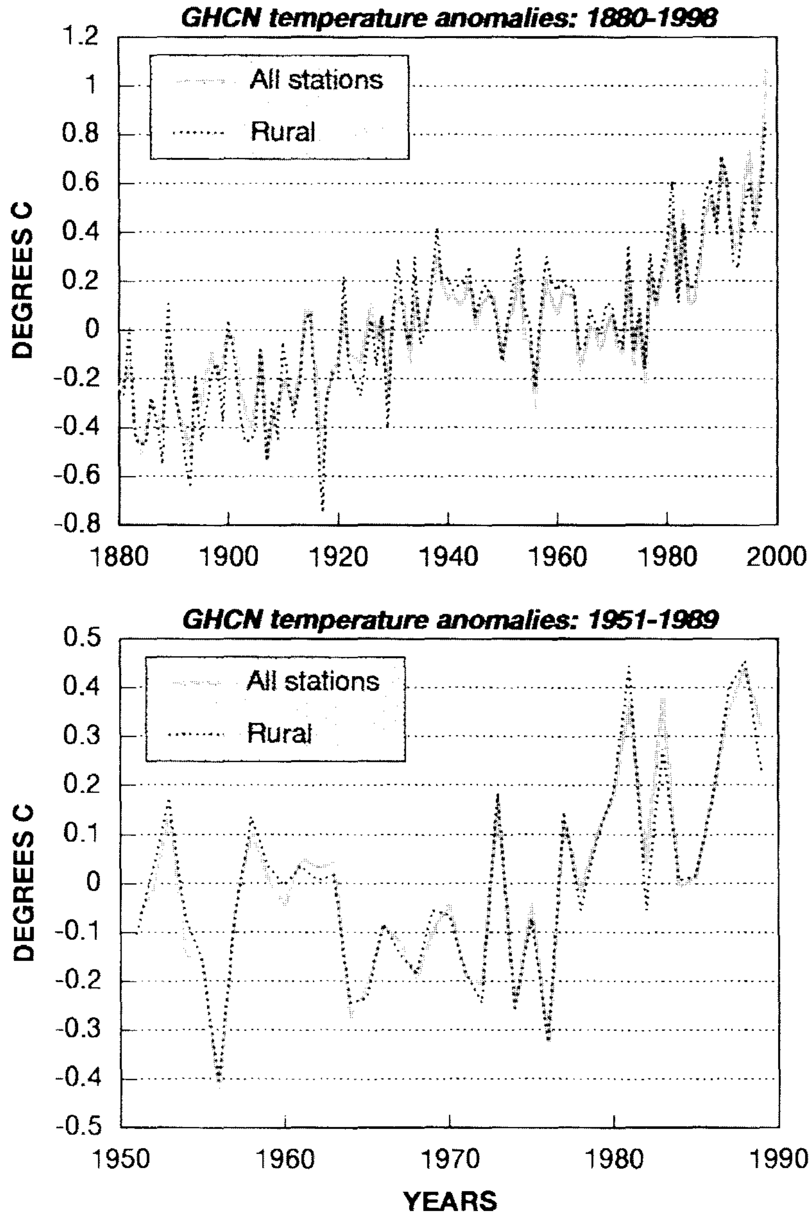


DMSP Nighttime Data

Satellite data from an instrument measuring visible light (DMSP OLS) was used to detect night time lights across the globe; including the U.S.. The data was used to identify rural stations that are unlikely to be affected by urban heat islands.

Source: Tom Karl, National Climatic Data Center, NOAA-NESDIS From section 4.1.3

Figure 9

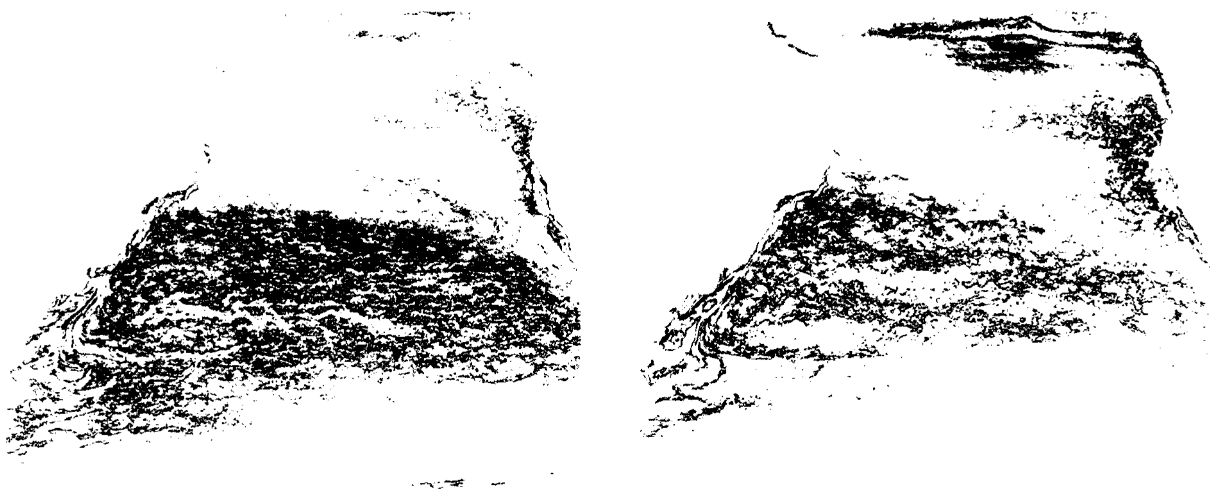


Urban Warming

Nighttime lights (satellite, Figure 8) and navigation charts define rural/non-rural area.

Source: Tom Karl, National Climatic Data Center, NOAA-NESDIS From section 4.1.3

Figure 10



(a) (b)

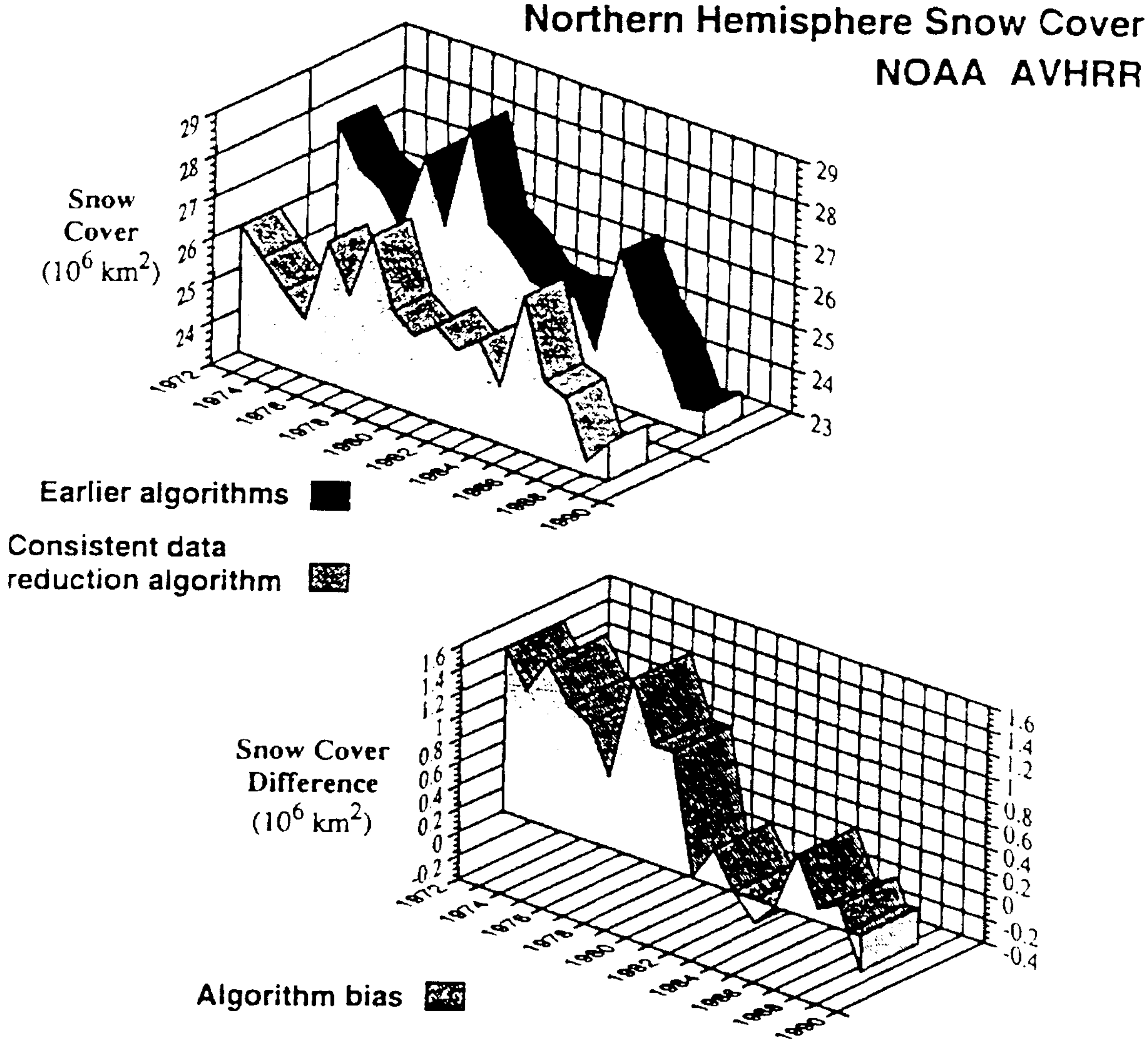
January **July**

Ocean Front Density (1985-1996)

This figure is an image of the South Atlantic using the AVHRR Sea Surface Temperature Multi-image edge detection algorithm on the 11-year time series. Red and yellow indicate colder temperatures while blues and purple indicate colder temperatures.

Source: Peter Cornillon, University of Rhode Island From section 4.1.4

Figure 11

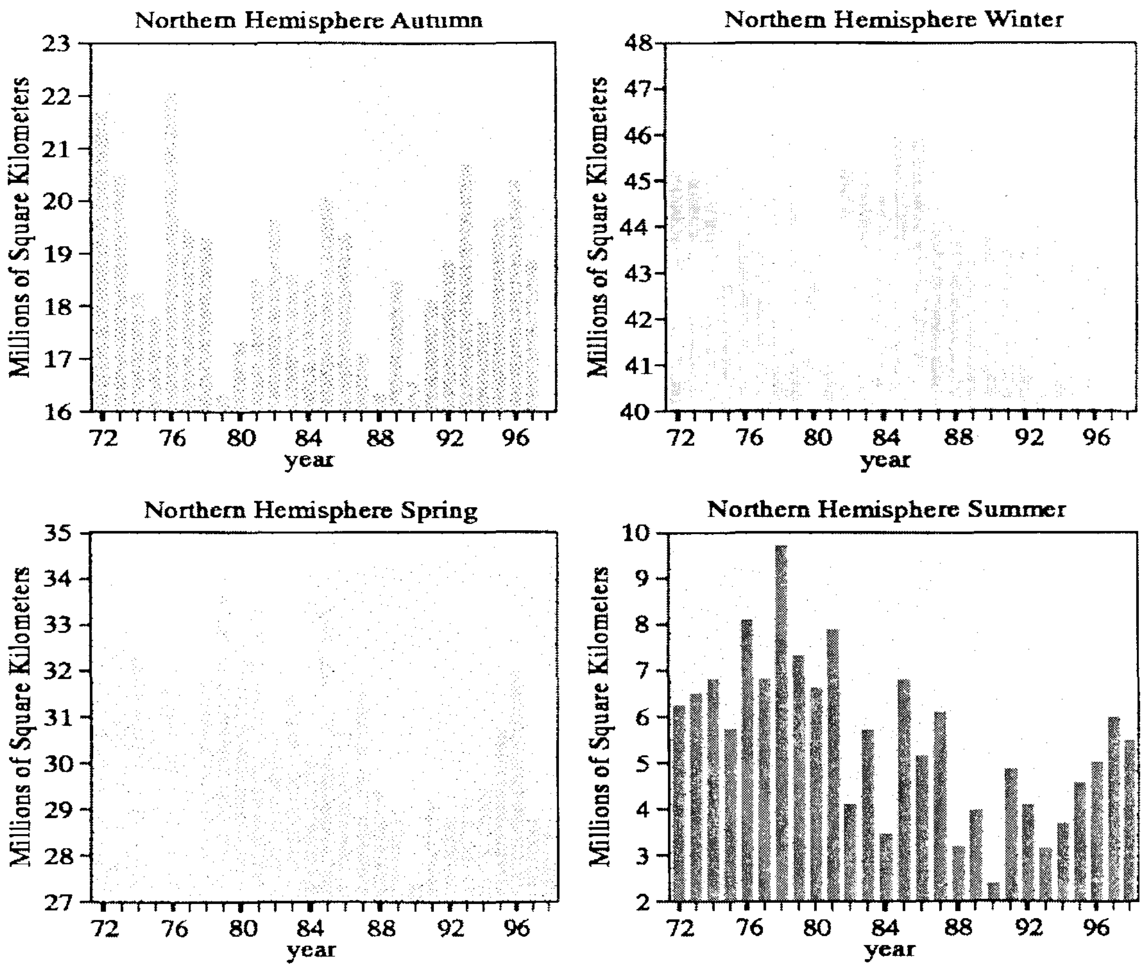


Northern Hemisphere Snow Cover, Algorithm Comparison

Changes in Northern Hemisphere snow cover as reported in IPCC (1990) and IPCC (1992) and the difference between the two time series.

Source: "Long-Term Climate Monitoring by the Global Climate Observing System" edited by Thomas R. Karl, p.67
From section 4.1.5

Figure 12

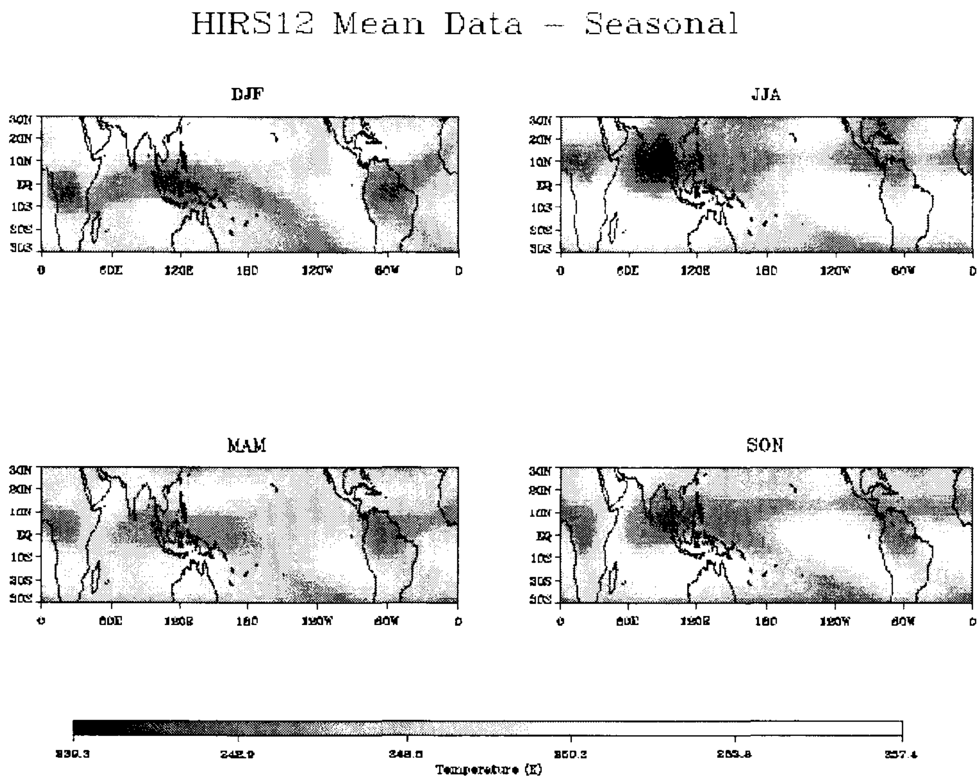


Snow Cover, 1973-Present

Satellite data on snow cover trends published in the 1990 IPPC report using the corrected data.

Source: Tom Karl, National Climatic Data Center, NOAA-NESDIS From section 4.1.5

Figure 13

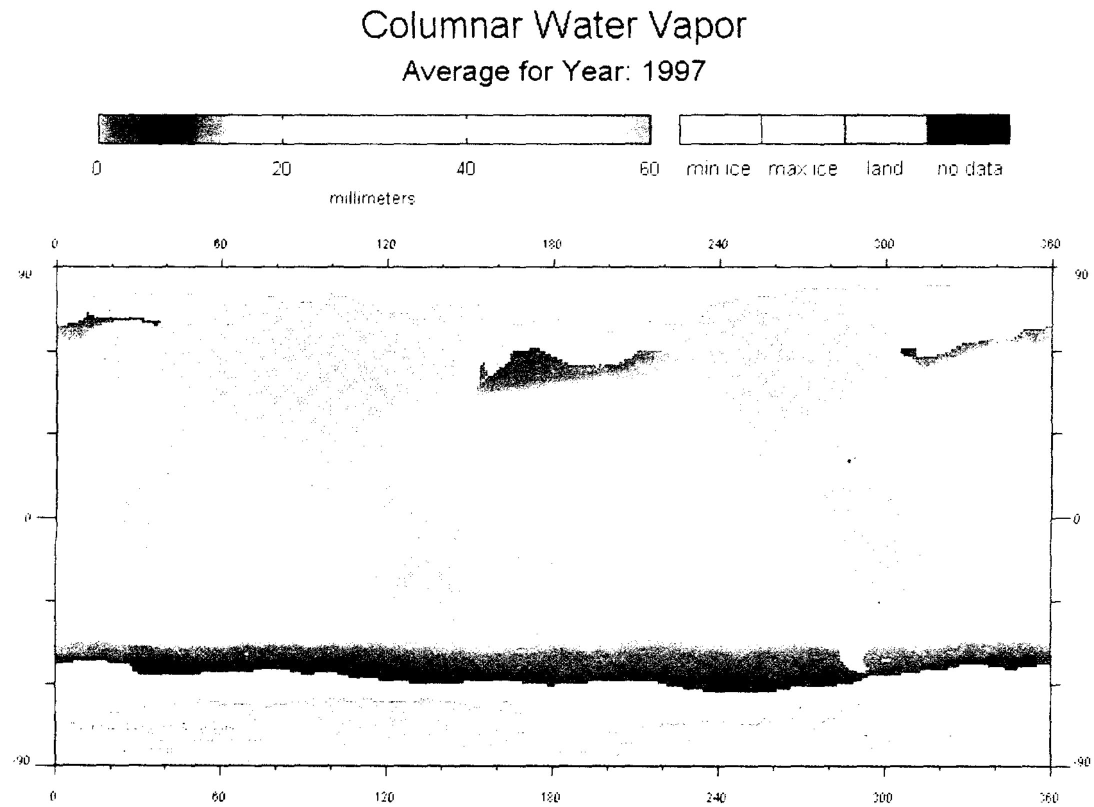


Upper Tropospheric Humidity Cycle

An analysis of the annual UTH cycle spanning a time period of 1979 through 1995 using inter-satellite calibrated HIRS data.

Source: John Bates, Climate Diagnostic Center, NOAA From section 4.2.3

Figure 14

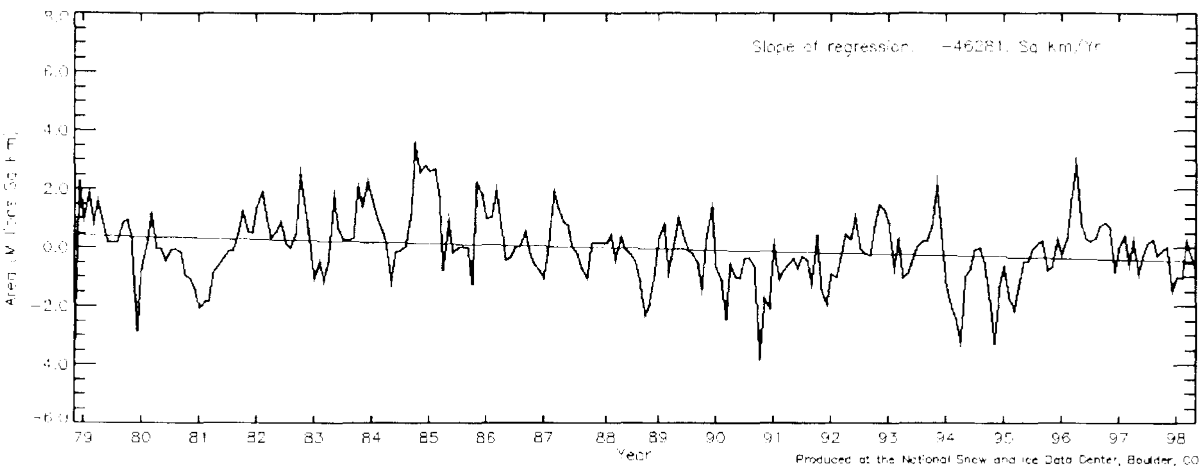


Columnar Water Vapor

An example of SSM/I water vapor data.

Source: Mike Goodman, Global Hydrology Resource Center. From section 4.2.5
Produced by Frank Wentz of Remote Sensing Systems, Santa Rosa, CA under contract to NASA, NASW-5038.

Figure 15

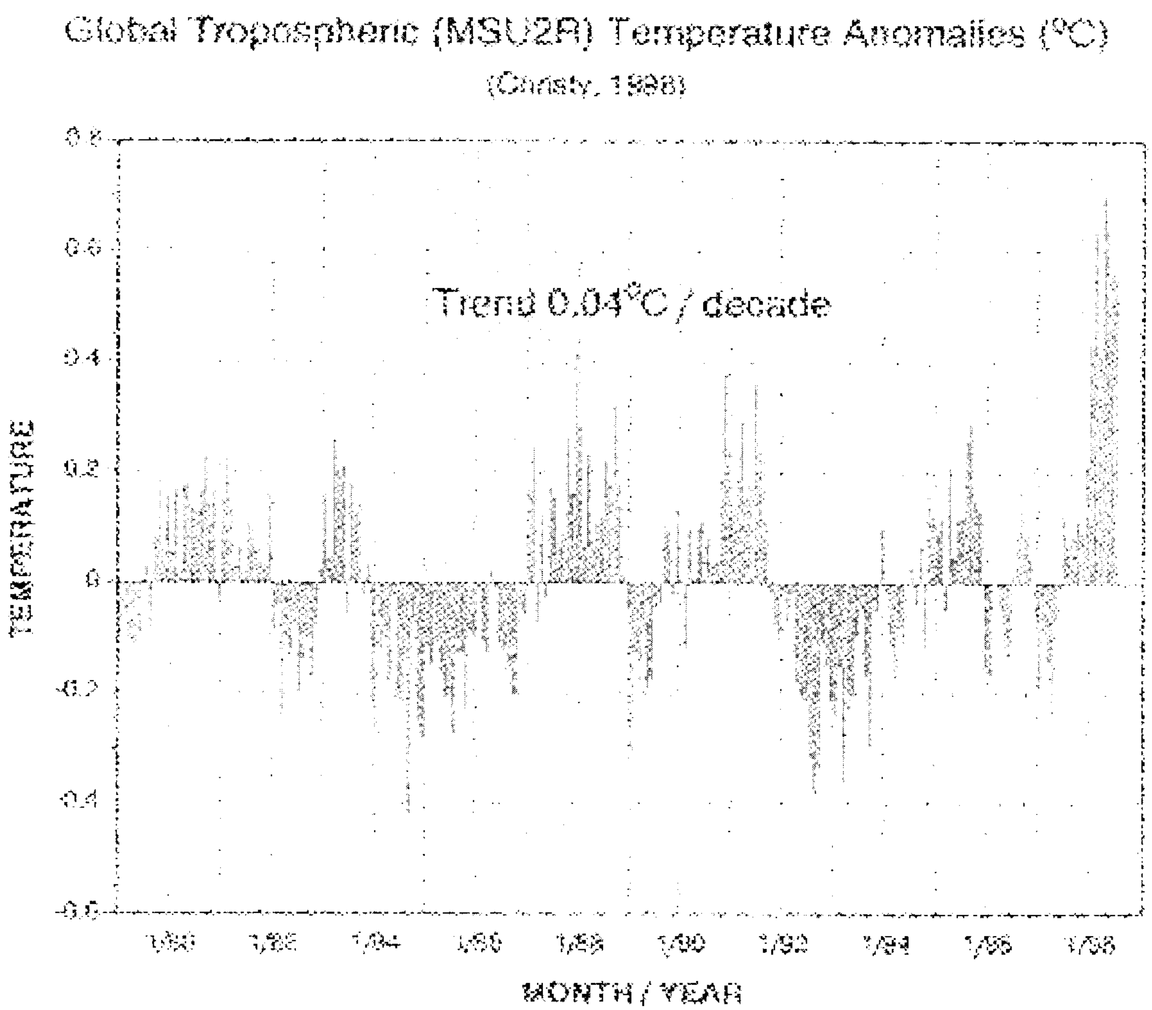


SSM/I and SMMR Snow Cover Trends

Passive microwave-derived (SMMR and SSM/I) snow covered area (x10^6 km^2) departures from monthly means for the Northern Hemisphere (1978-1998).

Source: Richard Armstrong and Mary Jo Brodzik, University of Colorado From section 4.2.5

Figure 16

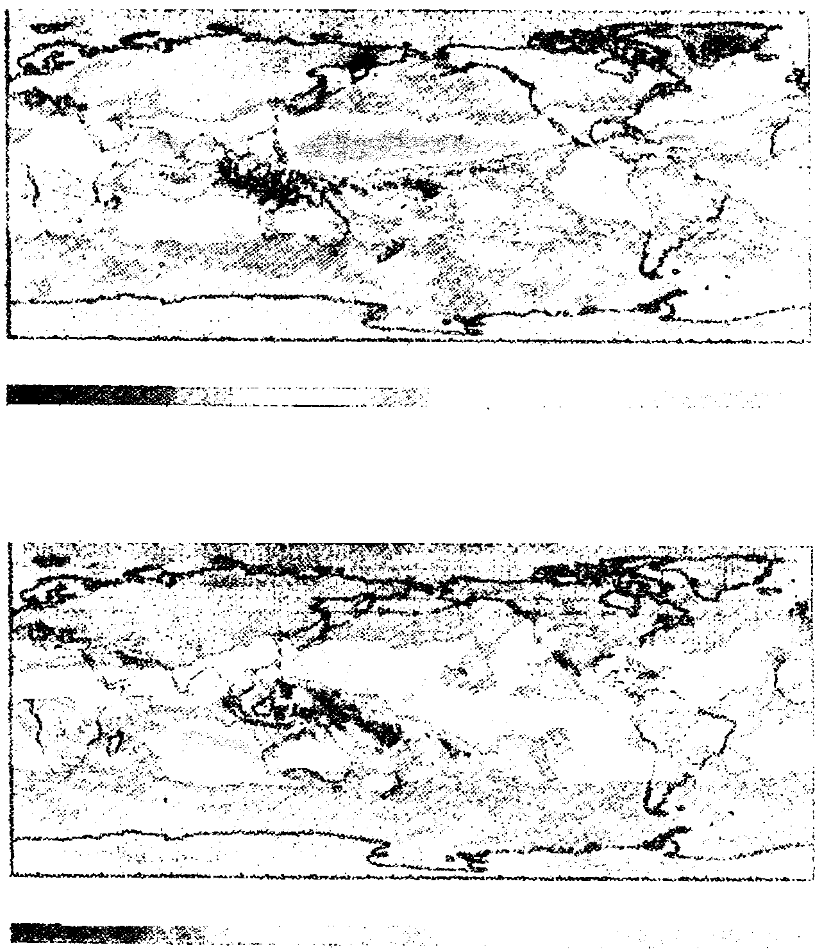


Temperature Changes in the Troposphere since 1979

Global temperature anomalies determined after the algorithms for computing the temperatures, correcting for contamination and instrument calibration have been improved.

Source: Mike Burgin, National Climatic Data Center, NOAA From section 4.4.1

Figure 17

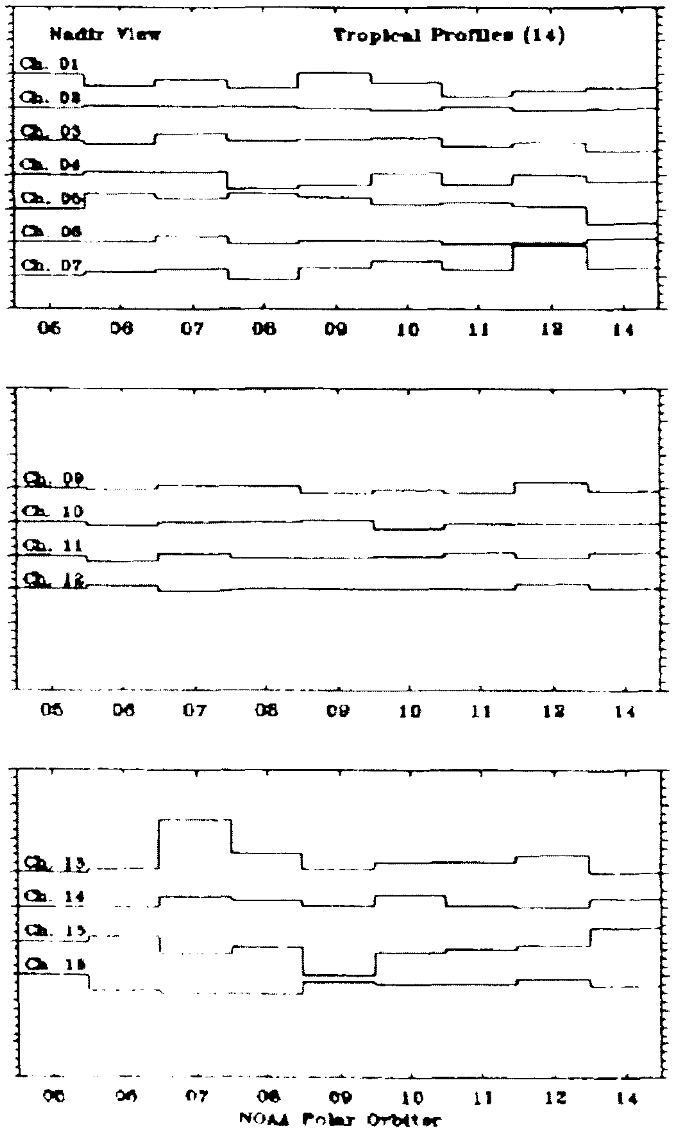


ENSO Extremes in the Upper Tropospheric Water Vapor

The difference between the amount of water vapor in El Niño, 1982-83 (upper panel) and La Niña, 1988-89 (lower panel) years revealed by analysis of HIRS data. The red end of the spectrum indicates high amounts of water vapor and the blue end of the spectrum indicates low amounts of water vapor. Note that the El Niño panel shows zonal bands of convection and drying and an intense hydrological cycle while the La Niña panel shows meridional bands of convection and drying and a weak hydrological cycle.

Source: Tom Karl, National Climatic Data Center, NOAA-NESDIS From section 4.4.3

Figure 18

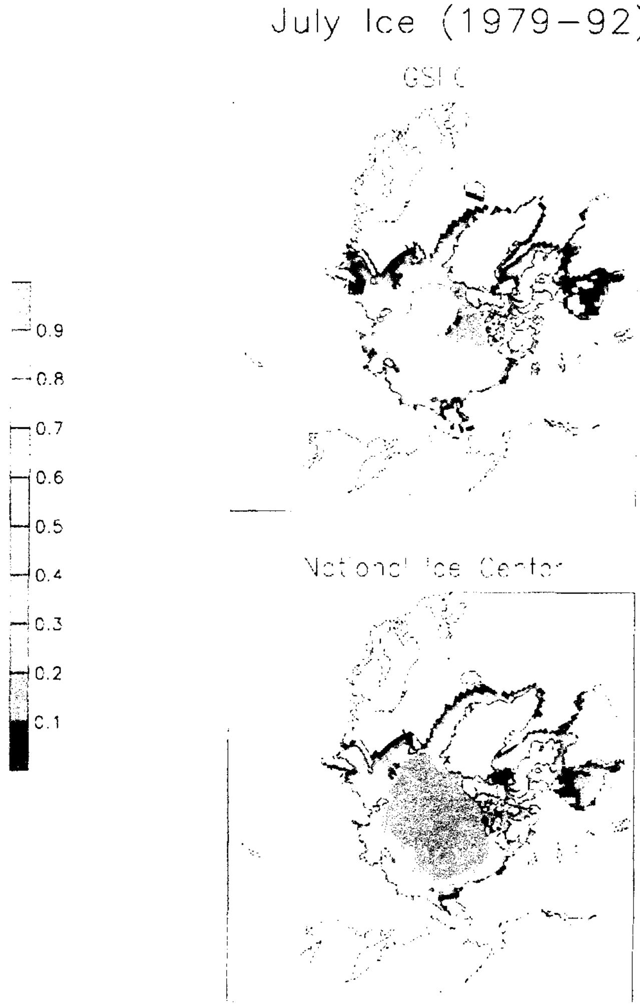


Intercalibration of HIRS Instruments

Calculated HIRS satellite-dependant brightness temperature biases. Among the corrections being made to HIRS data, potential systematic biases have been identified due to differences in the filters used in successive HIRS instruments.

Source: Tom Karl, National Climatic Data Center, NOAA-NESDIS From section 4.4.3

Figure 19



July Ice Cover (1979-1992)

The top panel shows data on the concentration of sea ice taken directly from satellites, which indicates nearly open water during the summer months around the north pole. But aircraft overflights and other data show that the sea ice has not melted, and it has been determined that the satellite instruments are sensing melt ponds that form on top of ice surfaces during the summer. A combined product of station, aircraft, and satellite data produced by the National Ice Center is shown in the lower panel.

Source: Tom Karl, National Climatic Data Center, NOAA-NESDIS From section 4.4.4

Appendix A

Workshop Attendees

Richard Armstrong	NASA / NOAA National Snow and Ice Data Center
John Bates	NOAA / ERL Climate Diagnostics Center
Francis Bretherton	University of Wisconsin, Space Science and Engineering Center
Josef Cihlar	Canadian Center for Remote Sensing
Margarita Conkright	NOAA - NESDIS National Oceanographic Data Center
Peter Cornillon	University of Rhode Island
Richard Cram	NOAA - NESDIS National Climatic Data Center
Pavel Groisman	UCAR / National Climatic Data Center
Wayne Faas	NOAA - NESDIS National Climatic Data Center
Dave Goodrich	NOAA / USGCRP
David Hofmann	NOAA Climate Monitoring and Diagnostics Laboratory
Greg Hunolt	SGT, Inc.
Bisher Imam	University of Arizona
Roy Jenne	National Center for Atmospheric Research
Tom Karl	NOAA - NESDIS National Climatic Data Center
Jack Kaye	NASA Head Quarters
Herb Kroehl	NOAA - NESDIS National Geophysical Data Center
Mike Loughridge	NOAA - NESDIS National Geophysical Data Center
Gwynneth Martin	GTOS Representative
Mark McCloy	NOAA - NESDIS
Martha Maiden	NASA Goddard Space Flight Center
Mark Parsons	NASA / NOAA National Snow and Ice Data Center
H. K. Ramapryian	NASA Goddard Space Flight Center
Cora Randall	LASP / University of Colorado
Mike Richman	University of Oklahoma
Wayne Rohde	USGS EROS Data Center
Steve Running	University of Montana
Matt Schwaller	NASA Goddard Space Flight Center
Richard Stolarski	NASA Goddard Space Flight Center
Kevin Trenberth	NCAR
Susan Zevin	NOAA - NESDIS
Victor Zlotnicki	NASA Jet Propulsion Laboratory

Appendix B

NASA / NOAA Memorandum of Understanding

The following is the text of the MOU signed by NASA and NOAA in July 1989. It is the basic agreement between the two agencies on transfer of the responsibility for active long-term archiving of NASA oceans and atmosphere data to NOAA as part of the overall cooperation between the agencies.

MEMORANDUM OF UNDERSTANDING
BETWEEN
THE NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
AND
THE NATIONAL OCEANIC AND ATMOPSHERIC ADMINISTRATION
FOR
EARTH OBSERVATIONS REMOTELY SENSED DATA PROCESSING,
DISTRIBUTION, ARCHIVING, AND RELATED SCIENCE SUPPORT

I. BACKGROUND

Since the 1960's, both the National Aeronautics and Space Administration (NASA) and the National Oceanic and Atmospheric Administration (NOAA) have observed, monitored, and studied the Earth's surface and environment from a variety of vantage points. As part of their overall mandates, both agencies play a major role in understanding the Earth's environment. NASA's enabling legislation (the National Aeronautics and Space Act of 1958 and the Supplemental Appropriations Act of 1962), and later amendments to this legislation, direct the agency to carry out a broad range of environmental observations, as well as space-based research, technology, monitoring, and other activities directed to understand the physics and chemistry of the upper atmosphere. Various statutes, including the Weather Service Organic Act, the Federal Aviation Act, the Land Remote Sensing Commercialization Act of 1984, the Coast and Geodetic Survey Act, and the Clean Air Act as amended, direct NOAA to make environmental observations; to monitor, understand, and predict climate conditions, and, as part of this mandate; to acquire, maintain and distribute long-term data bases, and to process and archive space-based data.

II. PURPOSE AND SCOPE OF THIS MEMORANDUM OF UNDERSTANDING

The purpose of this MOU is to establish the terms and conditions under which NASA and NOAA will cooperate as partners in the Earth Observing System (EOS). Such cooperation will encompass management of Earth observation remotely sensed data from spacecraft and aircraft, including processing of data, derived products, model results, distribution and archiving, and other related science activities and support.

III. NASA RESPONSIBILITIES

NASA will use its best efforts to:

1. Include, as appropriate, NOAA representatives in the planning, development, and implementation activities for experimental Earth observation data and information systems, including the EOS Data and Information System (EOSDIS).

2. Develop and implement, in coordination with NOAA, the United States Geological Survey (USGS), and others, the EOSDIS active short-term and long-term archives, which will include a distributed system providing data management, processing, access, distribution, and user interface capabilities for the EOS program. A major part of the EOSDIS will be located in a facility at Goddard Space Flight Center (GSFC). The active short and long-term archives for data from the EOS program will be placed in the EOSDIS facilities at GSFC and other facilities to be determined under terms of this and other MOUs.

3. Provide funding for the development, implementation, and operation of the active short-term archives, and appropriate science support activities, for the experimental remotely sensed data from the EOS program and other current and future experimental systems as agreed.

4. Transfer to NOAA, at a time to be determined, responsibility for active long-term archiving and appropriate science support activities for atmosphere and oceans data, as defined in accordance with approved coordinated program definition, development, and implementation activities and plans to be developed under this MOU (see V.2 and V.3).

5. Provide the necessary information and procedures to NOAA to permit NOAA to submit information about its holdings of data, including that from instruments an the NOAA Polar Orbiting Operational Environmental Satellites [POES(s)], the NOAA Geostationary Operational Environmental Satellites [GOES(s)], and European Polar-Orbiting Platforms [EPOP(s)], for entry into the EOSDIS Information Management Center (IMC), in a format to be agreed, in accordance with a plan to be developed under section V.2 and V.3.

6. Upon request and according to terms to be agreed, provide to NOAA other data, including data collected through programs conducted by NASA independently or with other domestic or foreign organizations, to support calibration analyses, system upgrades or other activities related to NOAA's operational mission. Such activities will be conducted within the constraints imposed by any external agreements.

7. When requested by NOAA, include in the EOSDIS contract(s) additional capabilities to support NOAA requirements at NOAA expense.

IV. NOAA RESPONSIBILITIES

NOAA will use its best efforts to:

1. Participate with NASA in planning, development, and implementation of experimental Earth observation data and information systems, including the EOSDIS.

2. Develop, implement, and manage an active archive for operational data, potentially collocated with a discipline related EOSDIS active archive, with access through a joint NASA/NOAA IMC, that will process and store data from NOAA POES(s), the operational meteorological payload on the EPOP(s), and future NOAA operational satellites and platforms.

3. Assume responsibility at a time to be agreed upon for active long-term archiving and appropriate science support activities for atmosphere and oceans data from the EOS program, as defined in accordance with approved coordinated program definition, development, and implementation plans to be developed under this MOU (see V.2 and V.3).

4. Provide information about NOAA's holdings of data, including that from instruments on the NOAA POES(s), GOES(s), and EPOP(s), for entry in the EOSDIS IMC, in a format to be agreed, in accordance with a plan to be developed under section V.2 and V.3.

5. Define at NOAA expense any additional capabilities which it wishes included at NOAA expense in EOSDIS contract(s) in a timely fashion, and so inform NASA.

V. JOINT RESPONSIBILITIES

NASA and NOAA will use their best efforts to:

1. Define an overall strategy for a cooperative program in Earth system science data management which identifies the important complementary roles of the EOS program and other experimental Earth observation systems, and NOAA's operational Earth observation and data and information systems.

2. Develop a joint plan for coordinated development of the short- and long-term archives, the IMC, and associated science support activities described in this MOU in accordance with a schedule to be agreed.

3. Prepare by an agreed date an initial Program Definition and Implementation Plan. The plan will identify and describe the scope of major elements covered by this agreement, including estimated funding requirements by each agency and implementation schedules.

4. Define a strategy for potentially collocating some part of NOAA's operational central data handling capability, data access and distribution system, and related functions and activities with an EOSDIS facility to be located at GSFC.

5. Ensure that catalog systems for NASA and NOAA data from the EOS program, including data from NOAA POES(s), GOES(s), and the operational data from EPOP(s), are interoperable.

6. Participate in joint presentations to NASA, DOC, NOAA, OMB, and the Congress, as necessary, to explain the essential role of each organization and funding needs for the elements of cooperation covered by this MOU.

7. Designate policy and technical points of contact in each agency for the implementation of this agreement.

VI. DATA ACCESS POLICY

1. Data from the NASA EOS prototype operational instruments will be made available to NOAA, in near real time and at no cost, for use in NOAA's operational mission.

2. Level 1b environmental data from NOAA instruments on NOAA POES(s) and EPOP(s) will be made available to NASA in near real-time and at no cost. Near real time access to other NOAA environmental data and products will be made available to NASA for its mission through arrangements to be agreed as part of the Program Definition and Implementation Plan (See V.2 and V.3). Access to and use of non-real-time archive data and products under NOAA's responsibility will be in accordance with legislative mandates.

3. Derived products from NASA EOS prototype operational instruments which result from NOAA operational activities may be used and distributed as part of NOAA's mission.

4. Data from NASA EOS program instruments, other than those from NASA EOS prototype operational instruments, will be made available to NOAA for use in its operational mission through arrangements to be agreed as part of the Program Definition and Implementation Plan (see V.2 and V.3). Access to and use of EOS research payload data by NOAA for research purposes will be the same as for other researchers from participating EOS entities.

VIII. FUNDING AND RESOURCES

1. NOAA and NASA will each be responsible for obtaining the resources for implementing their responsibilities under this MOU. Any commitment of resources is subject to the availability of appropriated funds.

2. Unless negotiated separately, joint and complementary activities between NOAA and NASA will be undertaken on a no-exchange-of-funds basis.

VIII. MANAGEMENT AND COORDINATION

1. If major changes are made in program schedules and progress that will affect the data management plans covered by this MOU, NASA and NOAA will consult and develop plans for appropriate action.

2. Both NASA and NOAA may freely release information regarding their own activities under this MOU; and may release information regarding the activities of the other agency after its approval. This item includes attribution of data sources by each party.

3. Responsibility for implementing the provisions of this MOU are assigned to the Director, Earth Science and Applications Division, NASA, and the Deputy Assistant Administrator for Satellite, Data, and Information Services, NOAA/NESDIS.

4. This MOU will enter into force as of the latter date of signature below. The agreement may be modified by mutual agreement of the parties and may be terminated by either party upon 180 days written notice. It will remain in force for the duration of the activities covered by the agreement.

For the National Oceanic and
And Atmospheric Administration

By: (signed) Tom Pyke

Date: 7 July 1989

Place: Washington, DC

For the National Aeronautics
and Space Administration

By: (signed) Len Fisk

Date: 21 July 1989

Place: Washington, DC

Appendix C

Long-Term Archiving - Science Requirements and Data Management

Beyond the USGCRP goals and objectives, the following are examples of science needs and requirements (from the NASA EOS program, the IPCC, and the Workshop on Long-Term Climate Monitoring) that can only be satisfied if Long-Term archiving of Earth science data is continued and extended to include EOS data and products and other key data sets such as NPOESS and NEXRAD. Also included here are derived data management requirements contained in the Data Management for Global Change Policy Statements and recommendations from the 1995 NRC study "Preserving Data on Our Physical Universe".

C.1 NASA's Earth Science Program

The Earth Observing System (EOS) is the cornerstone of the NASA Earth science program, NASA's contribution to the USGCRP.

The following are excerpts from "Science Strategy for the Earth Observing System" (Ghassem Asrar and Jeff Dozier, NASA Office of MTPE, July 1994):

> "Scientists need long-term, consistent measurements of the key physical variables that define the shifts in state and variability of Earth system components – the atmosphere, hydrosphere, cryosphere, oceans and land surface. Lacking these measurements, predictions of the complex responses of the Earth system to human activities and natural variations lack an adequate baseline to determine trends." (p. 13)

> "Determination of the magnitudes and spatial variations of global change requires consistent global measurements, over a long enough period to separate long-term trends from natural variability associated with seasons and other cyclical or periodic events. The duration of the observations must be at least 15 years, to span a solar sunspot cycle and several El Nino events. These observations must characterize the whole planet and its regional variations, and enable quantification of the processes that govern the Earth system." (p. 14)

> "To define the subtle, but critical, interactions between processes that modulate and regulate climate, researchers need long time series of data to differentiate secular trends from natural variability. For this reason, the EOS framework emphasizes continuous, long-term measurements." (p. 15)

> "The EOS science program will combine the measurements from the assembled suite of [EOS] instruments, using some instruments to correct the measurements of others or to provide data products derived from more than one instrument. This strategy permits the extraction of parameters that a single instrument cannot measure reliably. Global and regional data assimilation models that function as part of EOSDIS will use the

observations from the EOS platforms to provide continuous calculations of the principal fluxes between components." (p. 78)

C.2 International Panel on Climate Change

The following is taken from the report "Climate Change: The IPCC Assessment" (International Panel on Climate Change, WMO/UNEP, 1990):

> "It must be recognized that the requirements for climate monitoring are different from those for weather prediction. Failure to recognize this in the past means that there are a number of uncertainties which have been introduced into the long-term climatological time series... changes and improvements in observational networks should be introduced in a way which will lead to continuous, consistent long-term data sets sufficiently accurate to document changes and variations in climate. ... In order for these data [satellite observations] to be most useful, it is very important that they be analyzed and interpreted in conjunction with existing *in situ* data. High priority should be given to the blending or integration of space-based and *in situ* data sets in such a way as to build upon the strengths of each type of data. ... In planning new satellite observations for long-term monitoring, special attention should be paid to continuity of calibration and processing, archival, and access to the data." (p. 323, Chapter 11 "Narrowing the Uncertainties")

C.3 Recommendations from the Workshop on Long-Term Climate Monitoring Requirements

The Workshop on Long-Term Climate Monitoring Requirements was held at the National Climatic Data Center in January 1995. The results of the workshop were reported in the journal *Climatic Change* **31** (2-4), 1995 and published in a volume titled "Long-Term Climate Monitoring by the Global Climate Observing System" (Thomas Karl, Ed. Kluwer, 1996).

At the workshop, a series of technical papers were presented, each examining a different area of concern and the associated monitoring requirements (e.g. cryosphere, land cover changes, sea level changes, atmospheric circulation climate changes, etc.). Three workshop panels discussed climate forcings and feedbacks, climate responses and feedbacks and climate impacts. Each panel developed a set of recommendations which were discussed by the workshop participants. From this dialogue, a general set of recommendations was developed:

"1. Prior to implementing changes to existing systems or introducing new observing systems, an assessment on the effects on long-term climate monitoring should be standard practice.

2. Establish the capability for routinely assessing the quality of observations important for long-term climate monitoring, detection, and attribution. This should include the chain of activities involved in the processing of data into useful products.

3. Continued efforts should be made to improve the quality and volume of the historical data base, including long-term high resolution data capable of resolving important extreme climate events.

4. Improve existing and develop new information management systems related to the gathering, quality control, assembly, and archiving of climate data; and a system of telemetry, communication, and permanent storage dedicated to the objectives of the system.

5. Ensure that long-term climate assessments that focus on regular examination and re-examination of the climate data base are an integral part of GCOS. Up-to-date state-of-the-climate assessments should be produced.

6. Overlapping measurements of both the old and new observing systems for both in-situ and satellite data must become standard practice for critical climate variables.

7. Better satellite data and satellite product calibration is required such as: on-board calibration for all basic measurements of radiance, including the visible channels, and more appropriate 'ground truth' measurements with the sampling and physical characteristics as comparable as possible to the satellite products.

8. For polar orbiting satellites minimize the aliasing of the diurnal cycle of the parameter of interest, and most importantly ensure that diurnal sampling biases do not change with time by controlling the flight of the orbiting satellites, i.e. eliminate drift.

9. Pursue alternative means for calibrating stratospheric measurements from satellites. This includes better use of existing balloon borne systems, such as the ozonesonde network, and microwave measurements.

10. Ensure that the full and open international exchange of data is maintained at minimal cost of reproduction.

11. GCOS should encourage the development of metadata information systems that provide, among other types of information, factors that may result in time dependent biases. The system should be readily accessible. It is further recommended that metadata be given high priority and that GCOS make a similar recommendation to WMO."

The implications for long-term archive capabilities are apparent; long records with comprehensive coverage are necessary for the detection of subtle climate changes and trends. Two themes, the requirements for data continuity and for complete documentation are repeatedly emphasized throughout the volume:

1. Various aspects of climate monitoring will require use of several combinations of satellite and *in situ* measurements. Problems associated with comparison and

combination of these measurements underscores the need for careful and complete documentation of all the data sets involved;

2. Successful climate monitoring requires the use of long records. Most of these are obtainable only through combination of successive but different observing systems (both satellite and *in situ*). Appropriate intercomparisons can only be done properly if the archived data sets are well documented.

It is unlikely that future satellite or *in situ* observing systems can or will be engineered to more than partially mitigate intercomparison or continuity problems. Thus, the burden will fall on the archive capability to ensure the capture, preservation, and ready access of the information for all data sets so scientists will be able to use them for climate monitoring.

C.4 Data Management for Global Change Research – Policy Statements (July 1991)

These policy statements were written to facilitate full and open access to quality data for global change research. They were prepared in consonance with the goals of the U.S. Global Change Research Program and represent the U. S. Government's position on the access to global change research data.

- The Global Change Research Program requires an early and continuing commitment to the establishment, maintenance, validation, description, accessibility, and distribution of high-quality, long-term data sets.

- Full and open sharing of the full suite of global data sets for all global change researchers is a fundamental objective.

- Preservation of all data needed for long-term global change research is required. For each and every global change data parameter, there should be at least one explicitly designated archive. Procedures and criteria for setting priorities for data acquisition, retention, and purging should be developed by participating agencies, both nationally and internationally. A clearinghouse process should be established to prevent the purging and loss of important data sets.

- Data archives must include easily accessible information about the data holdings, including quality assessments, supporting ancillary information, and guidance and aids for locating and obtaining the data.

- National and international standards should be used to the greatest extent possible for media and for processing and communication of global data sets.

- Data should be provided at the lowest possible cost to global change researchers in the interest of full and open access to data. This cost should, as a first principle, be no more than the marginal cost of filling a specific user request. Agencies should act to streamline administrative arrangements for exchanging data among researchers.

- For those programs in which selected principal investigators have initial periods of exclusive data use, data should be made openly available as soon as they become widely useful. In each case, the funding agency should explicitly define the duration of any exclusive use period.

C.5 "Preserving Data on Our Physical Universe"

"Preserving Scientific Data on Our Physical Universe" reports on the study performed by the Steering Committee on the Long-Term Retention of Selected Scientific and Technical Records of the Federal Government, NRC, 1995. The study, funded by NARA and NOAA (with some NASA support), "identifies the major issues regarding efforts to archive and use data in the physical sciences, establishes retention criteria and appraisal guidelines for those data, reviews important technological advances and related opportunities, and proposes a new strategy to help ensure access to the data by future generations". The 'new strategy' was to be embodied in a National Scientific Information Resource (NSIR) Federation made up of independently funded and managed entities (including but not limited federal agencies) cooperating together in data management. The Global Change Data and Information System (GCDIS) effort was sited as an example of a prototype NSIR Federation. (A minority report offered that the elements of a federated approach already existed.)

The following are excerpts from the report:

> "A general problem prevalent among all scientific disciplines is the low priority attached to data management and preservation by most agencies. Experience indicates that new research projects tend to get much more attention than the handling of data from old ones, even though the payoff from optimal utilization of existing data may be greater."
>
> "The most important deficiencies are in the documentation, access, and long-term preservation of data in usable form. Insufficient documentation is a generic problem... few of the federal data centers can give adequate attention to long-term archiving because they are stretched thin by current demands and inadequate resources. Even the data that are archived may become inaccessible because they are not regularly migrated to new storage media as the hardware and software used to access the data become obsolete or inoperable."
>
> The study report, noting that NARA and other agencies must make decisions about which data to retain long-term and therefore need a basis for making those decisions contained recommendations regarding appraisal and retention criteria for physical science data:
>
> "As a general rule, all observational data that are non-redundant, useful, and documented well enough for most primary users should be permanently retained. ...For observational and experimental data, the following retention criteria should be used to determine whether a data set should be saved: uniqueness, adequacy of documentation (metadata), availability of hardware to read the data records, cost of replacement, and evaluation by

peer review. Complete metadata should define the content, format or representation, structure, and context of a data set. ...The appraisal on individual data sets ... should be performed by those most knowledgeable about the particular data – primarily the principal investigators and project managers. In some cases, they may need to involve an archivist or information resources professional to assist with issues of long-term retention."

Principles the committee believed should guide long-term retention of scientific and technical data include:

"The value of scientific data lies in their use. Meaningful access to data, therefore, merits as much attention as acquisition and preservation."

"Adequate explanatory documentation, or metadata, can eliminate one of today's greatest barriers to use of scientific data." "The problem of inadequate metadata is amplified when users are removed from the point of origin by being in a different discipline, by having a different level of expertise, or by time. Addressing this problem comprehensively will make data useful in the broadest possible context."

"The only effective and affordable archiving strategy is based on distributed archives managed by those most knowledgeable about the data," and a corollary "As a general principle, data collected by an agency should remain with that agency indefinitely." "Archive centers should generally be at the agencies or institutions that collect the data, and they should be responsible for archiving and providing access to the data as long as the agency's or institution's mission and scientific competence continue to encompass the subject field. Physical transfers of the data should be provided if possible, so agencies and institutions will need to allocate adequate resources to the entire life-cycle of their data holdings."

"Planning activities at the point of data origin must include long-term data management and archiving."

Appendix D

Acronym List

ADEOS-1 — Advanced Earth Observing System (Japan)

ASF — Alaska SAR Facility

ATMOS — Atmospheric Trace Model Spectrometer
Flown on four NASA Space Shuttle Missions in 1985, 1992, 1993, and 1994.

AVHRR — Advanced Very High Resolution Radiometer
A visible and infrared scanning radiometer flown aboard NOAA polar orbiting satellites since October, 1978.

CDC — Climate Diagnostics Center

CMDL — Climate Monitoring and Diagnostics Laboratory (NOAA)

COFUR — Cost of Filling the User Request
Meaning only those incremental costs associated with filling a request (media, shipping, handling, etc.) with no recovery of capital costs, data acquisition costs, etc.

DAAC — Distributed Active Archive Center
NASA's currently operational Earth science discipline-oriented data centers. They include: ASF (Polar Regions and SAR Data), EDC (Land Processes); GSFC (Upper Atmosphere, Global Biosphere, Atmospheric Dynamics, Geophysics), JPL (Physical Oceanography), LaRC (Radiation Budget, Tropospheric Chemistry, Clouds and Aerosols), NSIDC (Snow and Ice Cryosphere and Climate), ORNL (Biogeochemical Dynamics) and SEDAC (Human interactions in global environmental change and socioeconomic data).

DEM — Digital Elevation Model

DMSP — Defense Meteorological Satellite Program
Now de-classified DOD operational polar orbiting weather satellite series, flies instruments including the OLS and SSM/I.

DOC — Department of Commerce
Parent of NOAA

DOD — Department of Defense

DOE — Department of Energy

DOT — Department of Transportation
Parent of FAA

ECMWF — European Center for Medium Range Weather Forecasts

EDC EROS Data Center
In Sioux Falls, SD. Operated by the USGS; national data center for land surface remote sensing; holds large volume of aircraft; Landsat; declassified reconnaissance data. Agreed center for long-term archive of NASA EOS land remote sensing data. Hosts NASA's land surface DAAC.

EOS Earth Observing System

EOSDIS EOS Data and Information System

EOS-PM EOS-Passive Microwave

EPOP European Polar Orbiting Platforms

EROS Earth Resources Observation Systems

ERS European Remote Sensing Satellite
A set of two satellites, ERS-1 and ERS-2, launched by the European Space Agency in 1991 and 1995, respectively.

FAA Federal Aviation Administration

GAC Global Area Coverage
1 km resolution AVHRR data sampled on board the spacecraft to a nominal 4 km resolution

GCDIS Global Change Data and Information System

GCM General Circulation Model, or Global Climate Model

GLAS Geoscience Laser Altimeter System
Nadir-pointed laser altimeter to fly on IceSat mission in 2001.

GOES Geostationary Operational Environmental Satellite
Name of a series of NOAA operational satellites which collect visible and IR data from the Clarke belt orbit.

GOME Global Ozone Monitoring Experiment
Ozone instrument aboard ERS-2, launched in April 1995 and currently operational.

GSFC Goddard Space Flight Center (NASA)

HALOE Halogen Occultation Experiment
Flown on UARS, collecting data from October 1991 to the present.

HIRS High Resolution Infra-Red Sounder
Part of the TOVS package (along with MSU and SSU) flown aboard the NOAA series of polar orbiting satellites since late 1978.

ICESAT Ice Cloud & land Elevation Satellite

IMC Information Management Center of EOSDIS

IPCC International Panel on Climate Change.

IR Infra-red

ISCCP International Satellite Cloud Climatology Project
A WMO international project that collects data from all available geosynchronous satellites and AVHRR data from NOAA polar orbiters to produce cloud climatology products.

JPL Jet Propulsion Laboratory

LaRC Langley Research Center

LIMS Limb Infrared Monitor of the Stratosphere
Flown on Nimbus-7, collected data from October 1978 through May 1979.

LTA Long-Term Archive or Long-Term Archiving, as in LTA Program

MOU Memorandum of Understanding

MSU Microwave Sounding Unit
Passive microwave radiometer, part of the TOVS package (along with HIRS and SSU) flown aboard NOAA series polar orbiters.

NARA National Archives and Records Administration
Agency of the U.S. Government with jurisdiction of archives of Federal records (including NOAA and NASA data).

NASA National Aeronautics and Space Administration

NCAR National Center for Atmospheric Research
Located in Boulder CO

NCDC National Climatic Data Center
Located in Asheville NC, one of the three NOAA / NESDIS National Data Centers.

NCEP National Centers for Environmental Prediction
Formerly the National Meteorological Center.

NDVI Normalized Difference Vegetation Index
Vegetation extent mapping product determined from the visible channels of the NOAA AVHRR.

NESDIS National Environmental Satellite, Data, and Information Service
NOAA organization that operates NOAA and GOES series satellites, produces operational products, operates the NOAA National Data Centers, and where the NOAA Long-Term Archiving mission rests.

NEXRAD	Next Generation Weather Radar A neologism for the WSR-88D Doppler Radar now operating at 158 locations in the U.S. and a few locations internationally.
NGDC	National Geophysical Data Center Located in Boulder CO - one of the three NOAA / NESDIS National Data Centers.
NOAA	National Oceanic and Atmospheric Administration Located within DOC.
NODC	National Oceanographic Data Center Located in Washington DC - one of the three NOAA / NESDIS National Data Centers.
NODS	NASA Ocean Data System Predecessor at JPL to the Physical Oceanography DAAC.
NPOESS	National Polar Orbiting Environmental Satellite System Next generation of U.S. operational polar orbiting environmental satellites. Represents a convergence of the current NOAA series and the DMSP series.
NRC	National Research Council
NSF	National Science Foundation
NSIDC	National Snow and Ice Data Center Part of University of Colorado, Boulder and NOAA
NSIR	National Scientific Information Resource Foundation
NWS	National Weather Service
OLS	Operational Linescan System Instrument flown aboard the DMSP series of operational polar orbiters since 1976, back and forth sweeping rather than spinning mirror scanner, but still comparable to AVHRR. Noted for very low light level photo-multiplier visible data showing terrestrial and atmospheric light sources.
OMB	Office of Management and Budget
ORNL	Oak Ridge National Laboratory
PATMOS	Pathfinder Atmosphere Project A NOAA - NESDIS AVHRR reprocessing project that is part of the NOAA / NASA Pathfinder Program.
POES	Polar Orbiting Operational Environmental Satellite (NOAA)
SAGE	Stratospheric Aerosol and Gas Experiment
SAR	Synthetic Aperture Radar

Currently flying aboard Canadian (Radarsat), Japanese (JERS series), and European (ERS series) satellites.

SBUV — Solar Backscatter Ultra-Violet radiometer
Flown on Nimbus-7 and then on some NOAA polar orbiters as SBUV/2.

SDSD — Satellite Data Services Division
The group within NCDC responsible for the NOAA satellite archive.

SEDAC — Socioeconomic Data and Applications Center

SMMR — Scanning Multichannel Microwave Radiometer
Flown on Nimbus-7, provided data for the period 1978 - 1987.

SSEC — Space Science and Engineering Center
University of Wisconsin

SSM/I — Special Sensor Microwave / Imager
Flown on DMSP satellites since 1987.

SSM/IS — Special Sensor Microwave / Imager and Sounder
Successor to the SSM/I.

SST — Sea Surface Temperature

TBM — TeraBit Memory system
An Ampex system of 1977 vintage that recorded digital data on two inch-wide video tapes. Used as an on-line buffer and archive device by NOAA-NESDIS for the first five years of AVHRR and TOVS Level 1 data.

TIROS — Television InfraRed Observation Satellite

TOGA-COARE — Tropical Oceans and Global Atmosphere Program - Coupled Ocean-Atmosphere Response Experiment

TOMS — Total Ozone Mapping Spectrometer
Flown by NASA on Nimbus-7 in 1978, Meteor-3 in 1991, TOMS-Earth Probe in 1996, and ADEOS-1 in 1996. TOMS-Earth Probe is currently operating.

TOVS — TIROS Operational Vertical Sounder
Package of three instruments (HIRS, MSU and SSU) flown aboard NOAA operational polar orbiting satellites since 1978.

UARS — Upper Atmosphere Research Satellite (NASA)
Launched in October 1991.

USAF — U.S. Air Force

USGCRP — U. S. Global Change Research Program

Cooperative effort of U.S. agencies including NOAA and NASA, formed in 1991.

USGS — U. S. Geological Survey
Part of the U. S. Department of the Interior, parent organization of the EROS Data Center.

UTH — Upper Tropospheric Humidity

UV — Ultraviolet

VHRR — Very High Resolution Radiometer
Predecessor instrument to the AVHRR, flown on several NOAA operational polar orbiters prior to TIROS-N and NOAA-6.

VISSR — Visible Infra Red Spin Scan Radiometer
One of family of instruments flown aboard GOES and predecessor geosynchronous environmental satellites.

WMO — World Meteorological Organization

WSR-88D — Weather Surveillance Radar 1988 Doppler
A.k.a. NEXRAD, the new Doppler radar system deployed at 158 US and several foreign sites.

References

Asrar, G. and J. Dozier. 1994. Science Strategy for the Earth Observing System. NASA Office of MTPE.

Bates, J., J. X. Wu and D. L. Jackson. 1996. Interannual variability of upper-troposphere water vapor band brightness temperature. *Journal of Climate* **9**, 427-438.

Christy, J. R.. 1995. Temperature Above the Surface Layer. *Climatic Change* **31**, 455-474.

Elliott, W.. 1995. On Detecting Long-Term Changes in Atmospheric Moisture. *Climatic Change* **31**, 349-367.

Eyre J. R., G. A. Kelly, A. P. McNally, E. Andersson and A. Persson. 1993. Assimilation of TOVS radiance information through one-dimensional variational analysis. *Quarterly Journal of the Royal Meteorological Society* **119**, 1427-1463.

Farman, et al.. 1985. Large losses of total ozone in Antarctica reveal seasonal ClO_x/NO_x interaction. *Nature* **315**, 207.

Friis-Christensen, E. and L. Lassen. 1991. Length of the Solar Cycle: An Indicator of Solar Activity Closely Associated with Climate. *Science* **254**, 698-700.

Hansen, J., W. Rossow, B. Carlson, A. Lacis, L. Travis, A. Del Genio, I. Fung, B. Cairns, M. Mischenko and M. Sato. 1995 Low-Cost Long-Term Monitoring of Global Climate Forcings and Feedbacks. *Climatic Change* **31**, 247-271.

International Panel on Climate Change. 1990. Climate Change: The IPCC Assessment. WMO/UNEP.

Karl, Thomas R., Ed. 1996. Long-Term Climate Monitoring by the Global Climate Observing System. Kluwer Academic Publishers, The Netherlands. 518 pp.

Kidwell, K. 1997. NOAA Polar Orbiter Data Users' Guide. NOAA/NESDIS/NCDC.

McPeters, R. and G. Labow. 1996. An assessment of the accuracy of 14.5 years of Nimbus 7 TOMS version 7 ozone data by comparison with the Dobson network. *Geophysical Research Letters* **23**, 3695-3698.

Rusch et al.. 1998. *Journal of Geophysical Research* **103**, 8465-8475.

Spencer, R. W. and J. R. Christy. 1992a. Precision and Radiosonde Validation of Satellite Gridpoint Temperature Anomalies, Part I: MSU Channel 2. *Journal of Climatology* **5**, 847-857.

Spencer, R. W., and J. R. Christy. 1992b. Precision and Radiosonde Validation of Satellite Gridpoint Temperature Anomalies, Part II: A Tropospheric Retrieval and Trends During 1979-1990. *Journal of Climatology* **5**, 858-866.

Subcommittee on Global Change Research. 1998. Our Changing Planet: The FY99 U.S. Global Change Research Program. CENR.

Steering Committee on the Long-Term Retention of Selected Scientific and Technical Records of the Federal Government. 1995. Preserving Scientific Data on Our Physical Universe. NRC.

Stolarski, et al.. 1986. Nimbus-7 satellite measurements of the springtime Antarctic ozone decrease. *Nature* **322**, 808-811.

Stolarski, R. S., G. J. Labow and R. D. McPeters. 1997. Springtime Antarctic total ozone measurements in the early 1970s from the BUV instrument on Nimbus 4. *Geophysical Research Letters* **24** (5), 591-594.